Repurposing Bacterial CO_2-fixing Organelles Using Synthetic Engineering

Thesis submitted in accordance with the requirements of the

University of Liverpool for the degree of Doctor in Philosophy by

Yi Fang

April 2018

Department of Functional and comparative genomics, school of biological science,
Institute of integrative biology

Repurposing Bacterial CO2-fixing Organelles Using Synthetic Engineering

by Yi Fang

Institute of Integrative Biology

University of Liverpool

Dr. Luning Liu

Abstract

Bacterial microcompartments (BMCs) are protein-based organelles widespread among bacterial phyla and provide a means for compartmentalization of specific metabolic pathways. They sequester catalytic enzymes from the cytoplasm, using an icosahedral proteinaceous shell with selective permeability to metabolic molecules and substrates, to enhance metabolic efficiency. Carboxysomes were the first BMCs discovered, and their unprecedented capacity of CO_2 fixation allows cyanobacteria to make a significant contribution to global carbon fixation. The carboxysome was classified as α- and β-carboxysome according to the Rubisco form. The shell of the carboxysome ensures preferable transport of bicarbonate and CO_2 assimilation, making the carboxysome as an ideal nanobioreactor. Assembly studies and functional characterization of synthetic carboxysomes are important for understanding the assembly of carboxysomes in native and heterologous hosts. Moreover, there is an increasing interest in utilizing synthetic biology to construct synthetic carboxysomes or carboxysome shells in new hosts, i.e., higher plants, to enhance carbon fixation and productivity.

In Chapter 1, I summarized the research background and aims of the PhD project. In Chapter 3, I described the construction of a synthetic operon of β-carboxysomes from the cyanobacterium *Synechococcus elongatus* PCC7942 and the synthetic production of functional β-carboxysome structures in *Escherichia coli*. The structure, assembly, activity and interchangeability of synthetic β-carboxysomes were assessed using confocal, electron and atomic force microscopy, proteomics, immunoblot analysis, and enzymatic assays. To our knowledge, this is the first production of functional β-carboxysomes in heterologous organisms. In Chapter 4, the α-carboxysomes from the chemoautotroph *Halothiobacillus neapolitanus* were heterogeneous synthesized in *E. coli*. To improve the activity, the Rubisco activase CbbQO complex was introduced to the α-carboxysome expression system. We found that the addition of Rubisco activase CbbQ can improve the function of the synthetic carboxysomes and didn't affect the assembly of synthetic carboxysomes. In Chapter 5, the synthetic α-carboxysome shell was constructed and characterized. In Chapter 6, I summarized the results and prospect the future perspectives for my projects.

This study strengthens synthetic biology toolbox for generating not only functional carboxysome structures but also BMC-like organelles with tunable activities and BMC shell. It will inform the improvement of carboxysome engineering and the construction of functional CO_2-fixing modules for plant engineering as well as new nanoreactors.

Acknowledgments

It is a genuine pleasure to express my deep sense of thanks to my supervisor Dr. Luning Liu. His optimism and passion for science have inspired me. His scholarly advise, careful scrutiny, and scientific approach helped me to complete my research and dissertation and showed me spacious intellectual pursuit in the future.

I also have gratitude to my secondary supervisor Dr. Geraint Parry and Dr. Anthony Hall. I also want to thank Pro. Mark Caddick, Dr. Meriel Jones, and Dr. Heather Allison for their kind help and suggestions for my PhD study. Many thanks to the China Scholarship Council who supported me with a 4-year scholarship and the Institute of Integrative Biology who support me with a 6-month scholarship.

I thank all the members of the Liu Lab profusely for their company, lively intellectual discussions, and cooperation for my research. Especially thanks to Matthew Faulkner, who contributed all the Rubisco assays and the Atomic force microscopy experiments; Dr. Fang Huang, who helped me with Confocal microscopy; and Greg Dykes, who provided the Electron microscopy images. Also thanks to Dr. Selene Casella and Yaqi Sun, who accompany me for last four years, and Dr. Jorge Rodriguez-Ramos and Dr. Taiyu Chen, who gave me lots of suggestions. I also want show many thanks to Dr. James Johnson and Dr. Hannah McCue from Genemill, Dr. Deborah Simpson who provided proteomic analysis and our Co-operator group led by Prof Kerfeld, Berkeley and Prof PrIce, Australian National University.

I am indebted to my family, for their tireless supportduring my research period and all of my former supervisors: Dr. Geraint Parry, Dr. Chengchao Zheng and Dr. Changxiang Zhu, who encouraged me and were concerned for my research. Lastly, many thanks to my friends: Zimeng Zhang, Mengru Yang, Qiuyao Jiang, Dr. Pei Zhang, Dr. Yuefang Du, and Dr. Shancheng Cao.

Publications and Author's Contributions

The content Chapter 3 is based on the article:

Engineering and modulating functional cyanobacterial CO2-fixing organelles.
Fang, Y, Huang F, Faulkner M, Jiang QY, Dykes G, Yang MR, and Liu LN*.
Frontiers in Plant Science 9 (2018): 739. DOI: 10.3389/fpls.2018.00739.

I acknowledge the following collaborators for their contribution to the results described in these chapters. Unless specified below, all work was completed by the Author.

Dr. Fang Huang	Fluorescence microscopy Imaging
Matthew Faulkner	Rubisco Assay and data analysis Atomic force microscopy imaging and data analysis
Dr. Deborah Simpson	Proteomic analysis
Gregory Dykes	Transmission electron microscopy imaging
Prof PrIce	Provision of the CsoS1DCbbQ plasmid

Other publications produced during my PhD studies:

Light modulates the biosynthesis and organization of cyanobacterial carbon fixation machinery through photosynthetic electron flow.
Sun Y, Casella S, Fang Y, Huang F, Faulkner M, Barrett S, Liu LN*.
Plant Physiology, 2016, 171(1): 530-541. DOI:10.1104/pp.16.00107.

I helped to construct some plasmids for this article.

Contents

List of Figures

List of Tables

List of Abbreviations

2-PG	2-phosphoglycolate
3-PGA	3-phosphoglycerate
aa	Amino Acid
AFM	Atomic Force Microscopy
BMC	Bacterial microcompartment
BMC-H	BMC-hexameric protein
BMC-T	BMC-trimeric protein
BMC-P	BMC-pentameric protein
bp	base pair
CA	carbonic anhydrase
CBB	Calvin-Benson-Bassham
CCM	CO2-concentrating mechanisms
CFP	Cyan Fluorescent Protein
Ci	inorganic carbon
CO_2	carbon dioxide
cso	carboxysome operon
E.coli	*Escherichia coli*
eGFP	enhanced Green Fluorescent Protein
EP	encapsulation peptide
EUT	Ethanolamine utilization microcompartments
eYFP	enhanced Yellow Fluorescent Protein
Flp	flippase
FRT	Flippase recognition target
HCO_3^-	bicarbonate
H. neapolitanus	*Halothiobacillus neapolitanus*
IPTG	isopropyl-β-D-1 -thiogalactopyranoside

LB	Lysogeny Broth
LSUs	large subunits
O2	oxygen
OD600	Optical Density at 600 nm
PAGE	Polyacrylamide Gel Electrophoresis
PDU	1, 2-propanediol utilization microcompartments
Rca	Rubisco activase
RuBP	Ribulose 1,5-bisphosphate
Rubsico	Ribulose 1,5-bisphosphate carboxylase/oxygenase
SDS	Sodium Dodecyl Sulfate
sfGFP	superfolder Green Fluorescent Protein
SSUs	small subunits
Syn7942	*Synechococcus elongatus* PCC7942
TEM	Transmission Electron Microscopy
α-CcmK2	anti-CcmK2
α-RbcL	anti-RbcL

Chapter 1

Introduction

1.1 Bacterial microcompartments

Bacterial microcompartments (BMCs) are giant, polyhedral, proteinaceous organelles, distributed in at least 23 different bacterial species (Axen et al., 2014). All BMCs are consist of a icosahedral protein shell and sequestered enzymes, which can catalyze sequential reactions within these specialized constructions from the cytoplasm (Bobik et al., 2006; English et al., 1994; Kerfeld et al., 2010; Shively et al., 1973; Yeates et al., 2010; Yeates et al., 2011; Yeates et al., 2013). The outer shell contains approximately 4000 protein subunits (Yeates et al., 2008), and a further 10–15,000 protein molecules are encased inside (Cheng et al., 2008). The BMC are typically around 40~200nm in diameter (Aussignargues et al., 2015), and the combined mass of the macromolecular assembly is estimated to be between 100 and 600 MDa (Cheng et al., 2008; Frank et al., 2013). These sophisticated structures allow cells to confine chemical reaction in space and time, and enhance the efficiency of compartmentalized metabolic pathways (Frank et al., 2013).

BMCs mainly include the anabolic carboxysomes for carbon fixation found in cyanobacteria and some chemoautotrophs and anoxygenic phototrophs (Figure 1-1A) (Shively et al., 1973), as well as a variety of catabolic metabolosomes from heterotrophic bacteria, such as the propanediol utilization microcompartment (PDU) (Figure 1-1B) (Bobik et al., 1999) and ethanolamine utilization microcompartment (EUT) (Figure 1-1C) (Kofoid et al., 1999; Kerfeld et al., 2010; Rae et al., 2013; Kerfeld et al., 2015; Bobik et al., 2015). Although the functions of distinct BMCs vary in their encapsulated enzymes, the specialized groups of shell proteins were phylogenetically

related (Frank et al., 2013; Kerfeld et al., 2015). The BMC shell structurally resembles an icosahedral viral capsid, with hexameric proteins (BMC-H) containing a single PF00936 domain (Kerfeld et al., 2005); trimeric proteins (BMC-T) containing two PF00936 domains (Klein et al., 2009); and pentameric proteins (BMC-P) containing one PF03319 domain (Tanaka et al., 2008).

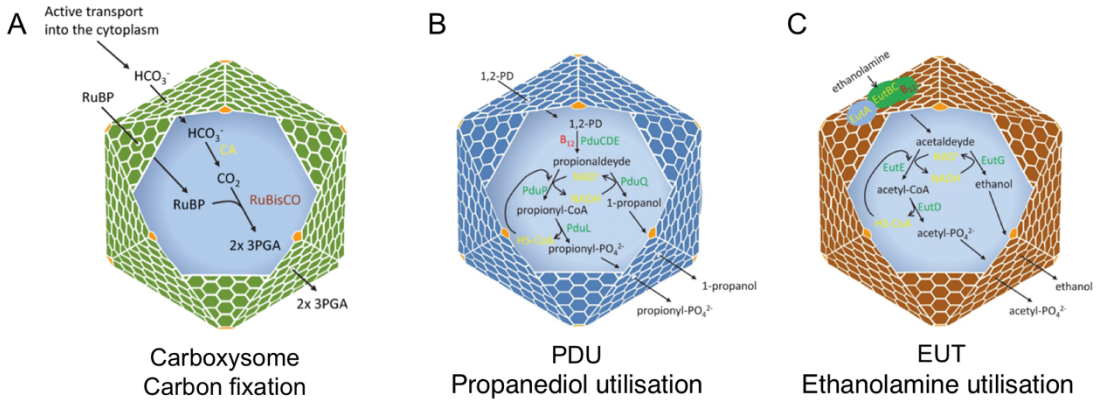

Figure 1-1. Models for Bacterial microcompartments (Bobik et al., 2015).
A. Diagram of CO_2 fixation in the carboxysome.
B. Diagram of propanediol metabolism in the PDU microcompartment.
C. Diagram of ethanolamine metabolism in the EUT microcompartment.

BMC-H proteins are the most abundant components in the shell and form the planar facets of the polyhedron (Cannon et al., 1983; Havemann et al., 2003; Roberts et al., 2012). The characteristic structures of BMC-H shell proteins from both carboxysomes and metabolosomes have been well studied (Kerfeld et al., 2005; Pang et al., 2014; Tsai et al., 2007; Crowley et al., 2008; Tanaka et al., 2008; Sagermann et al., 2009; Heldt et al., 2009; Crowley et al., 2010; Pang et al., 2011; Thompson et al., 2014; Sinha et al., 2014). The crystal structure showed a 6-fold symmetric hexamer containing a central pore (coincident with the cyclic symmetry axis) which is assumed to be constitutively open (Takenoya et al., 2010; Kerfeld et al., 2005; Tsai et al., 2007;

Tanaka et al., 2008; Tanaka et al., 2010). The size of these pores can admit the passage of smaller metabolites (e.g., bicarbonate into carboxysomes; alcohols and carboxylic acids across metabolosome shells) (Kerfeld et al., 2015), and the residues surrounding the pore tend to be polar and well conserved among orthologous shell proteins (Kerfeld et al., 2005; Tsai et al., 2007; Kinney et al., 2011; Cai et al., 2013; Sutter et al., 2013; Cai et al., 2014). Mutation of some of the residues leads to permeability defects (Sinha et al., 2014).

The BMC-P proteins assemblies a 5-fold symmetric pentamer and were deduced to cap the vertices of icosahedral shells (Figure 1-2) (Tanaka et al., 2008; Sutter et al., 2013; Kinney et al., 2011; Yeates et al., 2010). The crystal structures of BMC-T proteins indicate that they form trimers with pseudohexameric configuration and sometimes stacked into a double layer (Sagermann et al., 2009; Heldt et al., 2009; Klein et al., 2009; Cai et al., 2013; Larsson et al., 2017). The diameter of the BMC-T pore can be changed as a gate by the alternating conformation of the surrounding residues (Klein et al., 2009; Heldt et al., 2009; Cai et al., 2013; Takenoya et al., 2010; Pang et al., 2012). The open form of the pores in BMC-T proteins is typically larger than those in BMC-H shell proteins, likely facilitating the diffusion of the bulkier carboxysome metabolites, e.g., Ribulose 1,5-bisphosphate (RuBP) (Figure 1-2). Closure of the pores might be required to maintain the diffusive barrier and selectivity of the shell (Kerfeld et al., 2015, Klein et al., 2009).

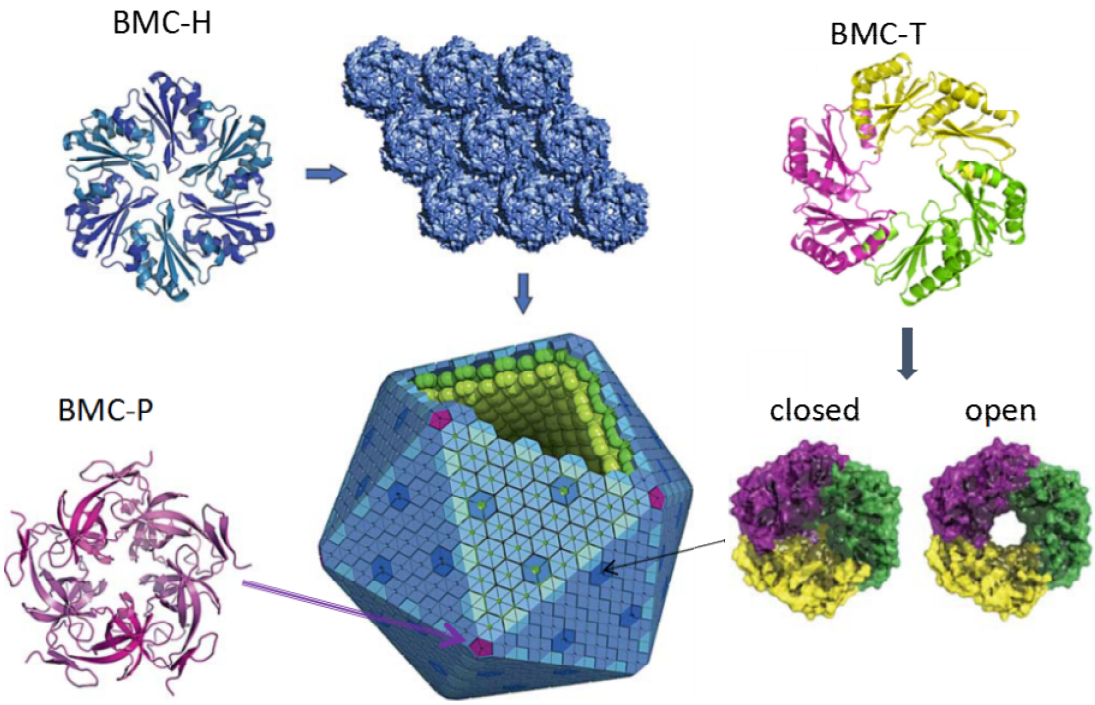

Figure 1-2. A model for the assembly of the BMC shell (Yeates et al., 2013). BMC-H (PDB ID: 2A1B, Kerfeld et al., 2005) proteins tile to make a molecular layer comprising the facets of the shell. BMC-P (PDB ID: 2QW7, Tanaka et al., 2008) proteins occupy vertex positions in the polyhedral shell, and BMC-T (PDB ID: 3F56, Klein et al., 2009) proteins can form open and closed pore structure and incorporate into the BMC-H facets.

The shell segregates metabolic core of the BMC from the cytosol of the cell, and it displays a remarkable capability to self-assemble (Kerfeld et al., 2015). The crystallization of a BMC-H protein exhibited that the hexamers tiled into uniformly oriented layers, evidently suggesting a model for facet formation (Kerfeld et al., 2005). All the PF00936 domains in BMC-H and BMC-T proteins contain a highly conserved motif (D/N-X-X-X-K), and it was shown that the antiparallel association of the lysine residues between neighboring hexamers was crucial for shell assembly (Klein et al., 2009; Pang et al., 2014; Sinha et al., 2014). In some crystal structures, BMC-H and BMC-T proteins tile into strips or layers (Kerfeld et al., 2005; Tanaka et al., 2008; Pang et al., 2014; Crowley et al., 2010). They can also form the fibrous tube, lattices, and

rolled up protein sheets *in vivo* when heterologously overexpressed (Parsons et al., 2010; Pang et al., 2014; Pitts et al., 2012; Heldt et al., 2009; Parsons et al., 2008). The purified BMC-H proteins under electron micrographs displayed ordered arrays, and under Atomic Force Microscopy (AFM), the proteins were observed to dissociate from and incorporate into assembled uniformly oriented shell layers *in vitro* (Cai et al., 2013; Lassila et al., 2014; Sutter et al., 2016). Overall, these results demonstrate the propensity of BMC shell proteins to self-assemble into highly ordered structures (Kerfeld et al., 2015; Sutter et al., 2017).

Understanding how thousands of proteins constitute the cargo and the shell, and then assemble and organize into functional organelles is essential for synthetic biology to design micro-bioreactors (Kerfeld et al., 2015). It was believed that the α-helical peptide regions could be the potential linker to recruit enzymes into BMC (Parsons et al., 2010; Fan et al., 2011; Fan et al., 2010; Kinney et al., 2012), like the N-terminal sequence extension of PduP (Fan et al., 2010) and the C-terminal of the carboxysomal protein CcmN (Kinney et al., 2012). However, the explicit mechanism of how to target and incorporate the cargos into the BMC shells has not been fully clarified.

BMCs have the potential capability for the improvement of cellular activity in the biosynthesis of non-native metabolites, such as ethanol and biodiesel for biofuel production (Jordan et al., 2016). The selectivity of the transport of the metabolites can be furnished by modulation of the shell component and pore properties. Based on the selective permeability and self-assembly of shell proteins, the BMC shells can be

used as nanomaterials, such as scaffolds or selectively permeable protein membranes. The designed selective permeable empty BMC shells could also be used as transport or storage capsules for drugs (Kerfeld et al., 2015, Frank et al., 2013, Lassila et al., 2014).

1.2 CO_2-concentration mechanisms as compensation for the imperfect enzyme Rubisco

Ribulose 1, 5-bisphosphate carboxylase/oxygenase (Rubisco) is the essential enzyme in the Calvin-Benson-Bassham (CBB) cycle of photosynthesis and is widely distributed in green plants, autotrophic bacteria, and algae. It catalyzes the first step of CO_2 fixation with ribulose 1, 5-bisphosphate (RuBP) to produce two molecules of 3-phosphoglycerate (3-PGA) (Figure 1-3) (Ellis et al., 1979).

There are four phylogenetically diverse clades of Rubisco that have been found in nature (Watson et al., 1999; Tabita et al., 2007). Form 1 Rubisco is the most abundant and has been most studied of the four forms. The structure of Form 1 Rubisco is a hexadecameric (L8S8) enzyme composed of eight large subunits (LSUs) and eight small subunits (SSUs) assembled into a symmetric structure (Schneider et al., 1992). The large subunits form into four dimers that then form a catalytic core (L8), with four small subunits on top and four at the bottom of this core. Plants and the majority of cyanobacteria utilise Form 1B Rubisco, whereas proteobacteria and a subgroup of marine cyanobacteria utilise Form 1A Rubisco (Schneider et al., 1992).

Rubisco is an imperfect enzyme regarding its enzymatic properties (Tcherkez et al., 2006). Rubisco can not only fix CO_2 but also fix O_2, and it has a poor capability of distinguishing CO_2 and O_2 (Figure 1-3). The Rubisco oxygenase reaction, known as photorespiration, produces one molecule 3-Phosphoglyceric acid (3-PGA) and one molecule 2-phosphoglycolate (2-PG). As a result, O_2 inhibits the efficiency of carbon fixation, and cells have to handle the toxic 2-PG compound (Anderson et al., 1971). The ecological and cellular environment also influence the function of Rubisco. In the history of the Earth's atmospheric evolutionary, the low level of O_2 didn't affect carbon fixation when Rubisco first evolves in cyanobacteria. However, the dramatic decrease in CO_2 concentration and increase in O_2 concentration in the atmosphere 400 to 350 million years ago led to O_2 becoming a major competitor for Rubisco carbon fixation (Shih et al., 2016; Badger et al., 2003; Whitney et al., 2011; Price 3t al., 2011d; Berner, 1990).

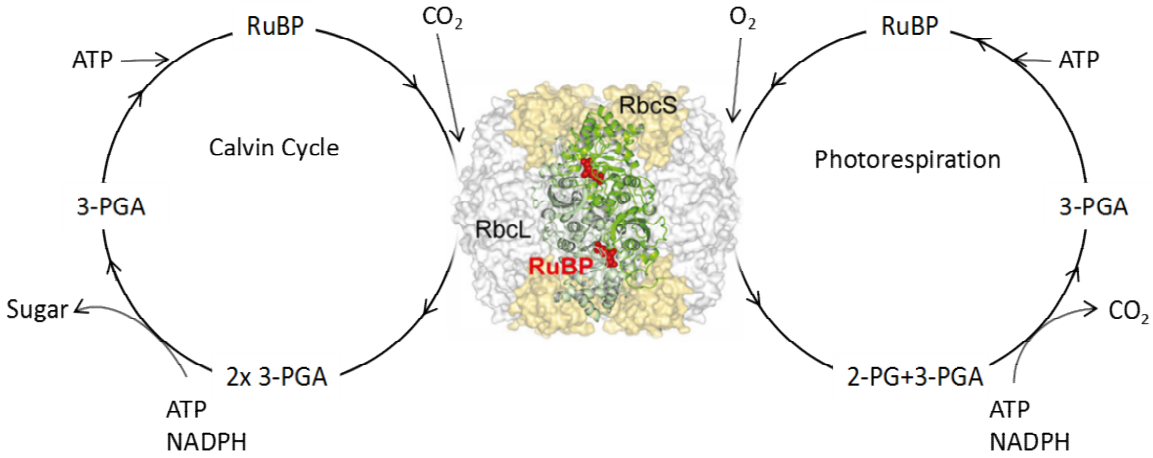

Figure 1-3. Competition between the Calvin-Benson cycle (photosynthesis) and photorespiration around RuBisCO (Rubisco structure from PDB: 1RCX, Bhat et al., 2017; Taylor et al., 1997).

For improving the efficiency of CO_2 fixation, one strategy exploited in most of the C_3 crop plants was to increase the Rubisco enzymatic specificity of CO_2 and improve the discrimination between CO_2 and O_2 (Badger et al., 1998; Tcherkez et al., 2006). However, in contrast, the carboxylation rate of Rubisco dropped more than three-fold in C_3 crop plants, which means the plants have to maximize the diffusive conductivity of CO_2 (Price et al., 2011d) and use more Rubisco enzymes (Price et al., 2012). Another strategy, found in cyanobacteria, many algae and some land plants (C_4 and CAM plants), was to develop active CO_2-concentrating mechanisms (CCM) to increase the CO_2 supply to Rubisco (Badger et al., 1998). Most of the important crop plants belong to C_3 species, which lack any CCM. Introducing functional CCM in these species to enhance photosynthesis has been considered as the promising strategy for increasing crop yield and water use efficiency (von Caemmerer et al., 2010; Price et al., 2011d; Parry et al., 2012).

The C_4 plants developed the cell specialized CCM. In mesophyll cells CO_2 is converted into a 4-carbon organic acid (C_4), and in bundle sheath cells C_4 acid is decarboxylated to release CO_2 that is then fixed by Rubisco (Wang et al., 2012; Hanson et al., 2016; Samal et al., 2013). The CAM plants developed a temporal separation of CO_2, which is initially fixed into C_4 during the night while Rubisco fixes CO_2 during the daytime. The CCM from C_4 plant engineering into the C_3 plant is studying by some researchers, but the division of the cells in C_4 plant means the key enzymes must localize properly in different cells (Hibberd et al., 2008; Sage, 2011; von Caemmerer et al., 2012). Cyanobacteria have evolved a different type of CCM

which is single cell-based, involving bicarbonate transporters, CO_2 uptake systems and subcellular compartmental carboxysomes (Gonzalez-Esquer et al., 2016; Price et al., 2011a; Gaudana et al., 2015). The powerful CCM elevates the CO_2 around Rubisco up to 1000-fold concentrated higher than in air (Badger et al., 2003). Engineering cyanobacterial CCM into C_3 plants might provide the promise for improving CO_2 concentration around Rubisco at the individual leaf chloroplast level, and as a result, we can increase the photosynthetic efficiency and crop yield (Price et al., 2007; McGrath et al., 2014).

According to the Rubisco phylogeny, cyanobacteria can be divided into two groups, the α-cyanobacteria with α-carboxysomes containing Form 1A Rubisco and β-cyanobacteria with β-carboxysomes encapsulating Form 1B Rbusico (Rae et al., 2013). Their CCM have similar designs and functions but with different protein components (Table 1-1) (Rae et al., 2013; Whitehead et al., 2014). The α-cyanobacteria survive in the ocean with high pH=8.2 and high dissolved carbon but low nutrient levels. Their cells are small and contain small genomes (1.6-2.8Mb), and they were thought to consume low energy and low flux, thus requiring a minimal CCM (Badger et al., 2006; Whitehead et al., 2014). In contrast, β-cyanobacteria are found in different kinds of habitats, like freshwater, estuarine, and hot springs (Badger et al., 2006). The cells are larger and possess larger genomes (2.2–3.6 Mb) than the α-cyanobacteria. Their CCMs are better studied and more complicated than those in α-cyanobacteria (Whitehead et al., 2014).

1.3 Ci Uptake

The first step of cyanobacterial CCM is to collect dissolved CO_2 and HCO_3^- as inorganic carbon (Ci) into the cells from aquatic environments. The β-cyanobacteria contains an array of carbon uptake transporters (Rae et al., 2011; Rae et al., 2013). Three types of bicarbonate transporters in the plasma membrane and two types of CO_2-uptake complexes located in the thylakoid membrane were identified in β-cyanobacteria. For bicarbonate transproters, BCT1 is a multimeric transporter belonging the traffic ATPase family with high binding affinity to HCO_3^- (Omata et al., 1999; Omata et al., 2002); SbtA is an inducible Na+-dependent HCO_3^- transporter, also with a high binding affinity to HCO_3^- (Shibata et al., 2002; Price et al., 2011c; Price et al., 2004); and BicA is a Na^+ -dependent transporter with a much lower affinity binding to HCO_3^-, belongs to SulP/SLC26 anion family but can provide higher rates of transport (Price et al., 2004; Price et al., 2011b; Shelden et al., 2010). The two CO_2-uptake complexes, NDH-I_4 and NDH-I_3, are derived from the NADHP dehydrogenase (NDH-I) complexes (Price et al., 2011a), which use NADPH as an electron donor to impel the conversion of CO_2 to HCO_3^- (Ohkawa et al., 2000; Folea et al., 2008; Xu et al., 2008; Price et al., 2002; Zhang et al., 2004). NDH-I_4 is constitutively expressed, while NDH-I_3 is induced under Ci limitation and has a higher affinity than NDH-I_4. They not only accumulate CO_2 diffused from outside of the cells but also CO_2 leaked from the carboxysome (Volokita et al., 1984; Badger et al., 1982; Espie et al., 1989; Miller et al., 1990; Fridlyand et al., 1996). Furthermore, because the charged HCO_3^- cannot easily diffuse across cell membranes like CO_2, the

conversion of CO_2 into HCO_3^- further concentrates Ci inside the cell (Zarzycki et al., 2012). Meanwhile, HCO_3^- is maintained because of the absence of carbonic anhydrase (CA) in the cytosol which can rapidly interconvert carbon dioxide and bicarbonate to maintain acid-base balance (Volokita et al., 1984; Price et al., 1989).

However, the Ci uptake is still poorly studied in α-cyanobacteria and was considered to contain a few constitutive Ci transporters (Rae et al., 2011; Beck et al., 2012). The homologous *sbtA2* and *bicA2* genes have been identified in α-cyanobacteria, but their roles remain unknown (Price et al., 2008; Badger et al., 2006; Rae et al., 2011; Shibata et al., 2002). BCT1 was only present in β-cyanobacteria (Badger et al., 2006) and an exceptional α-cyanobacterium, *Synechococcus* WH5701, which can survive in both freshwater and saltwater conditions. The CO_2 uptake system is significantly diverse between α-cyanobacteria species: some *Synechococcus* spp. species only possess the NDH-I_4 complex, *Prochlorococcus* species leak the entire system, whereas *Synechococcus* WH5701 and *Cyanobium* PCC 7001 possess both complexes (Price et al., 2007; Whitehead et al., 2014).

1.4 Carboxysomes

The carboxysome as the final step of cyanobacteria CCM, which encapsulates the Rubisco enzyme by a selectively permeable shell that acts as a physical barrier to permit the transition of HCO_3^-, RuBP, 3-PGA, and may block CO_2 efflux and O_2 influx (Cai et al., 2009; Price et al., 2007; Rae et al., 2013). The accumulated HCO_3^- by Ci

uptake can easily pass through the positively charged pores into the lumen of the carboxysome (Kerfeld et al., 2005; Tsai et al., 2007; Kinney et al., 2011), where the carboxysomal CA dehydrates HCO_3^- into CO_2 that is then fixed by Rubisco into 3-PGA. As a result, the CCMs of cyanobacteria accumulate in a much higher density of CO_2 around the encapsulated Rubisco and achieve a high catalytic rate for carboxylation.

The size of the α-carboxysome is approximately 100 nm, which is smaller than β-carboxysome that is around 150 nm (Faulkner et al., 2017). The genes to express α-carboxysome proteins tend to cluster in one operon, the *cso* operon (Roberts et al., 2012) whereas the genes of β-carboxysomes disperse into small operons in the genome containing *ccm* genes. The shell proteins of α- and β-carboxysomes, including BMC-H, BMC-T, and BMC-P proteins, have significant sequence homology regions and have been termed typical BMC domains as described before (Kerfeld et al., 2010; Rae et al., 2013). The CsoS1 from α-carboxysomes (Tsai et al., 2007; Cai et al., 2014) and CcmK from β-carboxysomes (Kerfeld et al., 2005, Tanaka et al., 2008; Cai et al., 2014; Samborska et al., 2012; Tanaka et al., 2009) form the major shell. CsoS4 from α-carboxysomes (Cai et al., 2009) and CcmL from β-carboxysomes (Keeling et al., 2014; Sutter et al., 2013) are BMC-P proteins that cap the vertices. CsoS1D from α-carboxysomes (Klein et al., 2009; Roberts et al., 2012) and CcmP, CcmO from β-carboxysomes (Cai et al., 2013; Marco et al., 1994) are BMC-T proteins that act as the gate of the shell (Table 1-1) (Kerfeld et al., 2016). Other carboxysome proteins may have similar function but do not carry sequence homology between the two types of carboxysomes, such as CsoS2 that is thought to be target Form 1A

Rubisco to the α-carboxysome shell (Cai et al., 2015) while CcmN works similarly in β-carboxysomes (Cameron, et al., 2013), without sharing any structure similarity. The CA encapsulated inside α- and β-carboxysomes are also different, with β-class CA CsoCA in α-carboxysome and γ-class CA CcmM and β-class CA CcaA in β-carboxysome (Kerfeld et al., 2016; Rae et al., 2013).

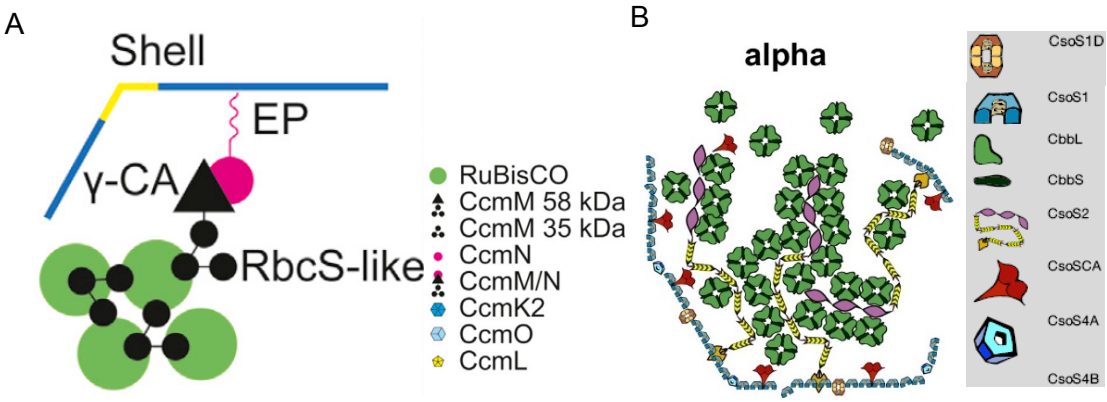

Figure 1-4. Model of β-carboxysome and α-carboxysome assembly.
A. Model of β-carboxysomes assembly, Rubisco (green) were aggregated by CcmM35 RbcS-like domain (black), and then recruit the shell proteins (CcmK2-blue and CcmL-yellow) by encapsulation peptide (EP) on CcmN (pink)(Cameron et al., 2013).
B. Model of α-carboxysome assembly, Rubisco (green) were captured inside of the shell by CsoS2 (yellow and pink) as well as the shell (blue) assembling (Kerfeld et al., 2016).

Furthermore, the Rubisco package in the two types of carboxysomes is also different from each other. The Rubsico enzymes inside β-carboxysomes were scaffolded in paracrystalline arrangement while the Rubisco inside α-carboxysomes are ordered layers bound to the shell (Rae et al., 2013; Kerfeld et al., 2016). Therefore, the assembly pathways of α-carboxysome and β-carboxysome were thought to be distinct. In α-carboxysomes, the partially assembled shells with attached Rubisco molecules were observed by Cryo-electron tomography, and the Rubsico knockout carboxysome

mutant resembled WT α-carboxysome structures, both suggesting that the shell and interior enzymes possess independent assemble pathways (Menon et al., 2008).The CsoS2 in α-carboxysomes from *Halothiobacillus neapolitanus,* which has have two isoforms because of the ribosomal frameshifting elements, was thought to be the essential factor to recruit Rubisco into carboxysomes (Figure 1-4B) (Chaijarasphong et al., 2016). The β-carboxysome appears to assemble from the inside core with soluble Rubisco nucleated into a procarboxysome by CcmM35 which contains three RbcS-like domains. The N-terminal of CcmN binds to the procarboxysome through interaction with the N-terminal γ-CA domain of CcmM58, and the C-terminus of CcmN form an amphipathic α-helix as the encapsulation peptide can help the procarboxysome interact with the assembling shell protein by CcmK and CcmO. And finally, the functional carboxysomes are fully encapsulated and released when CcmL is localized on the vertex (Figure 1-4A) (Cameron et al., 2013).

	β-carboxysomes		α-carboxysomes	
	Protein	Description	Protein	Description
Shell	CcmK2,3,4	Hexamer	CsoS1A,B,C	Hexamer
	CcmO,P	Trimer	CsoS1D	Trimer
	CcmL	Pentamer	CsoS4A,B	Pentamer
CA enzymes	CcaA	β-class CA	CsoS3	β-class
	CcmM	γ-class CA		
Structural proteins	CcmN	With encapsulation peptide	CsoS2	RuBisCO Nucleation protein
	CcmM	RuBisCO Nucleation protein		
Rubisco enzyme	RbcL	Form 1B Rubisco Large subunit	CbbL	Form 1A Rubisco large subunit
	RbcS	Form 1B Rubisco small subunit	CbbS	Form 1A Rubisco small subunit

Table 1-1. The model α-carboxysomal proteins from *Halothiobacillus neapolitanus* C2 and the model β-carboxysomal proteins from *Synechococcus elongatus* PCC 7942.

The locations of carboxysomes in cyanobacteria were indicated to be controlled by cytoskeletal filaments (Savage et al., 2010) which also help to control spacing and distribution of carboxysomes during cell division because the carboxysomes tend to be equally inheritable by daughter cells (Savage et al., 2010). However, another recent research showed that carboxysomes are self-organized along the nucleoid (MacCready et al., 2018), which leads to the regulation of the carboxysome position needs more further research. The number of the carboxysomes decreases strongly under high CO_2 concentration, and as it controls the important metabolic reaction of photosynthesis, the number and the spatial positioning of carboxysomes also

responds to light and the photosynthetic electron flow (Sun et al., 2006).

In addition to the carboxysomal elements, Rubisco activase (Rca) can structural remodel Rubisco into an active enzyme. Different forms of Rubisco need various types of Rca which belong to the AAA^+ family (ATPases Associated with diverse cellular Activities) (Neuwald et al., 1999). In β-cyanobacteria, Rca proteins were detected in *Synechococcus* sp. PCC 7942 (Friedberg et al., 1993) while in α-cyanobacteria CbbQO complex was determined as Rca (Tsai et al., 2015). The CbbQ protein encodes a ~30 kDa single AAA^+ subunits as the Rubisco activase and the CbbO was thought as an adapter between CbbQ and Rubisco (Tsai et al., 2015).

1.5 Engineering carboxysomes and cyanobacterial CCM

The α-carboxysomes from *H. neapolitanus* have been heterologously synthesized in *E. coli*, resulting in the production of well-formed structure, functional CO_2 fixation *in vitro* and *in vivo* (Bonacci et al., 2012). The α-carboxysome can also be functionally formed in the biotechnological platform species *Corynebacterium glutamicum* (Baumgart et al., 2017). These studies clarified that the α-carboxysome could self-assembly independently of the host which is useful for improving the engineering of the synthetic carboxysomes and further synthetic in the plant or other hosts. The intact β-carboxysomes haven't been reported to be heterologously synthesized. But the shell proteins of β-carboxysome from *Halothece* sp. PCC 7418 were heterogeneously expressed and formed empty polyhedral shells in *E. coli* (Cai et al.,

2016).

The benefits of engineering the cyanobacteria CCM into crop plant chloroplast could be significant. A stepwise strategy for engineering cyanobacterial CCM into C_3 plant chloroplasts has been proposed (Price et al., 2012). The first step is to introduce Ci transporter to the chloroplast envelope. Installing BicA and/or SbtA transporter into chloroplast inner envelope could increase 5%~15% of CO_2 fixation rates at constant substomatal CO_2 levels (Price et al., 2011d). The second step is to build functional carboxysomes in the chloroplast stroma. The third step is to remove CA from the stroma to complete the environment in the chloroplast that only contain carboxysomal- CA. And the final step is to build functional NDH-1 CO_2 uptake complexes in thylakoid membranes to reduce CO_2 leakage (Price et al., 2012).

Attempts have been made to install cyanobacterial CCM into plant chloroplasts. The expression of BicA from *Synechococcus* PCC 7002 in tobacco chloroplast was succeeded but didn't affect the photosynthetic CO_2 assimilation rates, which implies that installed BicA had no activity (Pengelly et al., 2014). The β-carboxysome proteins CcmK2, CcmM, CcmL, CcmO and CcmN from *Synechococcus elongatus* PCC 7942 were transiently transformed into tobacco chloroplasts and formed highly organized structures resembling empty microcompartments (Lin et al., 2014a). The tobacco Rubisco was also replaced by the cyanobacteria Rubisco with/without RbcX and CcmM35, but resulted in a slower growth than wild-type plants and required high CO_2 conditions (Lin et al., 2014b; Occhialini et al., 2016).

Based on the previous observations, it appears that more factors need to be considered in carboxysome engineering, such as the Rubisco chaperonin GroEL/ES which will facilitate Rubisco assembly (Hauser et al., 2015; Andersson et al., 2008), Rubisco activase which remodels functional Rubisco, and Rubisco assembly factor Raf1, which maintains environmentally regulated Rubisco homeostasis (Kolesinski et al., 2017). Therefore, engineering carboxysomes in *E. coli* are important to verify the essential components for future research.

1.6 Summary of this study

The carboxysomes are the final step of cyanobacterial CCM and play an essential role in bacterial CO_2 fixation. The remarkable ability of the carboxysome can increase the efficiency of CO_2 fixation has attracted the attention of researchers. The aim of my PhD study is to explore the assembly mechanisms and engineering strategies of carboxysomes in bacterial systems. There are three objectives. The first objective is to construct an entire β-carboxysome in *E. coli* (Chapter 3). By combining all the five operons of the β-carboxysome from *Synechococcus elongatus* PCC 7942 and expressing this the inducible plasmid in *E. coli*, I produced carboxysome-like structures in *E.coli* with Rubisco activity. In addition, I studied the interchangeability of BMC building blocks: the large subunit of Rubisco Form 1A can be incorporated into β-carboxysomes, the major shell protein CsoS1A from α-carboxysomes and PduA from PDU microcompartment can be recruited to the shell of β-carboxysomes. It gives

us the possibility to modify carboxysome composition and structure for new functions. The second project is to achieve a better functional synthetic α-carboxysomes by adding the Rubisco activase CbbQO complex (Chapter 4). Then I constructed and characterized the empty shells of α-carboxysomes in *E. coli* (Chapter 5). As a part of BMC with CO_2 accumulation, the carboxysome shell is an ideal nanocage for metabolized reaction.

1.7 References

Anderson, L. E. 1971. "Chloroplast and cytoplasmic enzymes. II. Pea leaf triose phosphate isomerases." *Biochim Biophys Acta* 235(1): 237-244.

Andersson, I. and A. Backlund. 2008. "Structure and function of Rubisco." *Plant Physiol Biochem* 46(3): 275-291.

Aussignargues, C., B. C. Paasch, R. Gonzalez-Esquer, O. Erbilgin and C. A. Kerfeld. 2015. "Bacterial microcompartment assembly: The key role of encapsulation peptides." *Commun Integr* Biol 8(3): e1039755.

Axen, S. D., O. Erbilgin and C. A. Kerfeld. 2014. "A taxonomy of bacterial microcompartment loci constructed by a novel scoring method." *PLoS Comput Biol* 10(10): e1003898.

Badger, M. R. and G. D. Price. 2003. "CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution." *J Exp Bot* 54(383): 609-622.

Badger, M. R. and T. J. Andrews. 1982. "Photosynthesis and Inorganic Carbon Usage by the Marine Cyanobacterium*, Synechococcus* sp." *Plant Physiol* 70(2): 517-523.

Badger, M. R., G. D. Price, B. M. Long and F. J. Woodger. 2006. "The environmental plasticity and ecological genomics of the cyanobacterial CO_2 concentrating mechanism." *J Exp Bot* 57(2): 249-265.

Badger, M. R., T. J. Andrews, S. Whitney, M. Ludwig, D. C. Yellowlees, W. Leggat and G. D. Price. 1998. "The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae." *Canadian Journal of Botany* 76(6): 1052-1071.

Baumgart, M., I. Huber, I. Abdollahzadeh, T. Gensch and J. Frunzke. 2017.

"Heterologous expression of the *Halothiobacillus neapolitanus* carboxysomal gene cluster in *Corynebacterium glutamicum*." *J Biotechnol* 258: 126-135.

Beck, C., H. Knoop, I. M. Axmann and R. Steuer. 2012. "The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms." BMC Genomics 13(1): 56.

Berner RA. 1990. “Atmospheric carbon dioxide levels over phanerozoictime.” *Science* 249: 1382–1386.

Bhat, J. Y., G. Thieulin-Pardo, F. U. Hartl and M. Hayer-Hartl. 2017. "Rubisco Activases: AAA+ Chaperones Adapted to Enzyme Repair." *Front Mol Biosci* 4: 20.

Bobik, T. A. 2006. "Polyhedral organelles compartmenting bacterial metabolic processes." *Appl Microbiol Biotechnol* 70(5): 517-525.

Bobik, T. A., B. P. Lehman and T. O. Yeates. 2015. "Bacterial microcompartments: widespread prokaryotic organelles for isolation and optimization of metabolic pathways." *Mol Microbiol* 98(2): 193-207.

Bobik, T. A., G. D. Havemann, R. J. Busch, D. S. Williams and H. C. Aldrich. 1999. "The propanediol utilization (pdu) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol degradation." *J Bacteriol* 181(19): 5967-5975.

Bonacci, W., P. K. Teng, B. Afonso, H. Niederholtmeyer, P. Grob, P. A. Silver and D. F. Savage. 2012. "Modularity of a carbon-fixing protein organelle." *Proc Natl Acad Sci U S A* 109(2): 478-483.

Cai, F., B. B. Menon, G. C. Cannon, K. J. Curry, J. M. Shively and S. Heinhorst. 2009. "The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier." *PLoS One* 4(10): e7521.

Cai, F., M. Sutter, J. C. Cameron, D. N. Stanley, J. N. Kinney and C. A. Kerfeld. 2013. "The structure of CcmP, a tandem bacterial microcompartment domain protein from the beta-carboxysome, forms a subcompartment within a microcompartment." *J Biol Chem* 288(22): 16055-16063.

Cai, F., M. Sutter, S. L. Bernstein, J. N. Kinney and C. A. Kerfeld. 2014. "Engineering bacterial microcompartment shells: chimeric shell proteins and chimeric carboxysome shells." *ACS synthetic biology* 4(4): 444-453.

Cai, F., S. L. Bernstein, S. C. Wilson and C. A. Kerfeld. 2016. "Production and Characterization of Synthetic Carboxysome Shells with Incorporated Luminal Proteins." *Plant Physiol* 170(3): 1868-1877.

Cai, F., Z. Dou, S. L. Bernstein, R. Leverenz, E. B. Williams, S. Heinhorst, J. Shively, G. C. Cannon and C. A. Kerfeld. 2015. "Advances in Understanding

Carboxysome Assembly in *Prochlorococcus* and *Synechococcus* Implicate CsoS2 as a Critical Component." *Life* 5(2): 1141-1171.

Cameron, J. C., S. C. Wilson, S. L. Bernstein and C. A. Kerfeld. 2013. "Biogenesis of a bacterial organelle: the carboxysome assembly pathway." *Cell* 155(5): 1131-1140.

Cannon, G. and J. Shively. 1983. "Characterization of a homogenous preparation of carboxysomes from *Thiobacillus neapolitanus*." *Archives of Microbiology* 134(1): 52-59.

Chaijarasphong, T., R. J. Nichols, K. E. Kortright, C. F. Nixon, P. K. Teng, L. M. Oltrogge and D. F. Savage. 2016. "Programmed Ribosomal Frameshifting Mediates Expression of the alpha-Carboxysome." *J Mol Biol* 428(1): 153-164.

Cheng, S., Y. Liu, C. S. Crowley, T. O. Yeates and T. A. Bobik. 2008. "Bacterial microcompartments: their properties and paradoxes." *Bioessays* 30(11-12): 1084-1095.

Crowley, C. S., D. Cascio, M. R. Sawaya, J. S. Kopstein, T. A. Bobik and T. O. Yeates. 2010. "Structural insight into the mechanisms of transport across the *Salmonella enterica* Pdu microcompartment shell." *Journal of Biological Chemistry* 285(48): 37838-37846.

Crowley, C. S., M. R. Sawaya, T. A. Bobik and T. O. Yeates. 2008. "Structure of the PduU shell protein from the Pdu microcompartment of *Salmonella*." *Structure* 16(9): 1324-1332.

Ellis, R. J. 1979. "The most abundant protein in the world." *Trends in biochemical sciences* 4(11): 241-244.

English, R. S., S. C. Lorbach, X. Qin and J. M. Shively. 1994. "Isolation and characterization of a carboxysome shell gene from *Thiobacillus neapolitanus*." *Mol Microbiol* 12(4): 647-654.

Espie, G. S., A. G. Miller and D. T. Canvin. 1989. "Selective and Reversible Inhibition of Active CO_2 Transport by Hydrogen Sulfide in a Cyanobacterium." *Plant Physiol* 91(1): 387-394.

Fan, C. and T. A. Bobik. 2011. "The N-terminal region of the medium subunit (PduD) packages adenosylcobalamin-dependent diol dehydratase (PduCDE) into the Pdu microcompartment." *J Bacteriol* 193(20): 5623-5628.

Fan, C., S. Cheng, Y. Liu, C. M. Escobar, C. S. Crowley, R. E. Jefferson, T. O. Yeates and T. A. Bobik. 2010. "Short N-terminal sequences package proteins into bacterial microcompartments." *Proc Natl Acad Sci U S A* 107(16): 7509-7514.

Faulkner, M., J. Rodriguez-Ramos, G. F. Dykes, S. V. Owen, S. Casella, D. M. Simpson, R. J. Beynon and L. N. Liu. 2017. "Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes." *Nanoscale* 9(30):

10662-10673.

Folea, I. M., P. Zhang, M. M. Nowaczyk, T. Ogawa, E.-M. Aro and E. J. Boekema. 2008. "Single particle analysis of thylakoid proteins from *Thermosynechococcus elongatus* and *Synechocystis* 6803: Localization of the CupA subunit of NDH-1." *FEBS letters* 582(2): 249-254.

Frank, S., A. D. Lawrence, M. B. Prentice and M. J. Warren. 2013. "Bacterial microcompartments moving into a synthetic biological world." *J Biotechnol* 163(2): 273-279.

Fridlyand, L., A. Kaplan and L. Reinhold. 1996. "Quantitative evaluation of the role of a putative CO_2-scavenging entity in the cyanobacterial CO_2-concentrating mechanism." *Biosystems* 37(3): 229-238.

Friedberg, D., K. Jager, M. Kessel, N. Silman and B. Bergman. 1993. "Rubisco but not Rubisco activase is clustered in the carboxysomes of the cyanobacterium *Synechococcus* sp. PCC 7942: Mud-induced carboxysomeless mutants." *Molecular Microbiology* 9(6): 1193-1201.

Gaudana, S. B., J. Zarzycki, V. K. Moparthi and C. A. Kerfeld. 2015. "Bioinformatic analysis of the distribution of inorganic carbon transporters and prospective targets for bioengineering to increase Ci uptake by cyanobacteria." *Photosynth Res* 126(1): 99-109.

Gonzalez-Esquer, C. R., S. E. Newnham and C. A. Kerfeld. 2016. "Bacterial microcompartments as metabolic modules for plant synthetic biology." *The Plant Journal* 87(1): 66-75.

Hanson, M. R., M. T. Lin, A. E. Carmo-Silva and M. A. Parry. 2016. "Towards engineering carboxysomes into C3 plants." *The Plant Journal* 87(1): 38-50.

Hauser, T., J. Y. Bhat, G. Milicic, P. Wendler, F. U. Hartl, A. Bracher and M. Hayer-Hartl. 2015. "Structure and mechanism of the Rubisco-assembly chaperone Raf1." *Nat Struct Mol Biol* 22(9): 720-728.

Havemann, G. D. and T. A. Bobik. 2003. "Protein content of polyhedral organelles involved in coenzyme B12-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar Typhimurium LT2." *J Bacteriol* 185(17): 5086-5095.

Heldt, D., S. Frank, A. Seyedarabi, D. Ladikis, J. B. Parsons, M. J. Warren and R. W. Pickersgill. 2009. "Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilization in *Clostridium kluyveri*." *Biochem J* 423(2): 199-207.

Hibberd, J. M., J. E. Sheehy and J. A. Langdale. 2008. "Using C_4 photosynthesis to increase the yield of rice-rationale and feasibility." *Current opinion in plant biology* 11(2): 228-231.

Jordan, P. C., D. P. Patterson, K. N. Saboda, E. J. Edwards, H. M. Miettinen, G. Basu,

M. C. Thielges and T. Douglas. 2016. "Self-assembling biomolecular catalysts for hydrogen production." *Nat Chem* 8(2): 179-185.

Keeling, T. J., B. Samborska, R. W. Demers and M. S. Kimber. 2014. "Interactions and structural variability of beta-carboxysomal shell protein CcmL." *Photosynth Res* 121(2-3): 125-133.

Kerfeld, C. A. and M. R. Melnicki. 2016. "Assembly, function and evolution of cyanobacterial carboxysomes." Current opinion in plant biology 31: 66-75.

Kerfeld, C. A. and O. Erbilgin. 2015. "Bacterial microcompartments and the modular construction of microbial metabolism." *Trends Microbiol* 23(1): 22-34.

Kerfeld, C. A., M. R. Sawaya, S. Tanaka, C. V. Nguyen, M. Phillips, M. Beeby and T. O. Yeates. 2005. "Protein structures forming the shell of primitive bacterial organelles." *Science* 309(5736): 936-938.

Kerfeld, C. A., S. Heinhorst and G. C. Cannon. 2010. "Bacterial microcompartments." *Annu Rev Microbiol* 64: 391-408.

Kinney, J. N., A. Salmeen, F. Cai and C. A. Kerfeld. 2012. "Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly." *J Biol Chem* 287(21): 17729-17736.

Kinney, J. N., S. D. Axen and C. A. Kerfeld. 2011. "Comparative analysis of carboxysome shell proteins." *Photosynth Res* 109(1-3): 21-32.

Klein, M. G., P. Zwart, S. C. Bagby, F. Cai, S. W. Chisholm, S. Heinhorst, G. C. Cannon and C. A. Kerfeld. 2009. "Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport." *J Mol Biol* 392(2): 319-333.

Kofoid, E., C. Rappleye, I. Stojiljkovic and J. Roth. 1999. "The 17-gene ethanolamine (eut) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins." *J Bacteriol* 181(17): 5317-5329.

Kolesinski, P., M. Rydzy and A. Szczepaniak. 2017. "Is RAF1 protein from *Synechocystis* sp. PCC 6803 really needed in the cyanobacterial Rubisco assembly process?" *Photosynthesis research* 132(2): 135-148.

Larsson, A. M., D. Hasse, K. Valegard and I. Andersson. 2017. "Crystal structures of beta-carboxysome shell protein CcmP: ligand binding correlates with the closed or open central pore." *J Exp Bot* 68(14): 3857-3867.

Lassila, J. K., S. L. Bernstein, J. N. Kinney, S. D. Axen and C. A. Kerfeld. 2014. "Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function." *Journal of molecular biology* 426(11): 2217-2228.

Lin, M. T., A. Occhialini, P. J. Andralojc, J. Devonshire, K. M. Hines, M. A. Parry and M. R. Hanson. 2014a. "β-Carboxysomal proteins assemble into highly organized

structures in *Nicotiana* chloroplasts." *The Plant Journal* 79(1): 1-12.

Lin, M. T., A. Occhialini, P. J. Andralojc, M. A. Parry and M. R. Hanson. 2014b. "A faster Rubisco with potential to increase photosynthesis in crops." *Nature* 513(7519): 547-550.

MacCready, J. S., P. Hakim, E. J. Young, L. Hu, J. Liu, K. W. Osteryoung, A. G. Vecchiarelli and D. C. Ducat. 201). "Protein Gradients on the Nucleoid Position the Carbon-fixing Organelles of Cyanobacteria." *bioRxiv* 334813.

Marco, E., I. Martinez, M. Ronen-Tarazi, M. I. Orus and A. Kaplan. 1994. "Inactivation of ccmO in *Synechococcus* sp. Strain PCC 7942 Results in a Mutant Requiring High Levels of CO_2." *Appl Environ Microbiol* 60(3): 1018-1020.

McGrath, J. M. and S. P. Long. 2014. "Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis." *Plant Physiology* 164(4): 2247-2261.

Menon, B. B., Z. Dou, S. Heinhorst, J. M. Shively and G. C. Cannon. 2008. "*Halothiobacillus neapolitanus* carboxysomes sequester heterologous and chimeric Rubisco species." *PLoS One* 3(10): e3570.

Miller, A. G., G. S. Espie and D. T. Canvin. 1990. "Physiological aspects of CO_2 and HCO_3^- transport by cyanobacteria: a review." *Canadian Journal of Botany* 68(6): 1291-1302.

Neuwald, A. F., L. Aravind, J. L. Spouge and E. V. Koonin. 1999. "AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes." *Genome research* 9(1): 27-43.

Occhialini, A., M. T. Lin, P. J. Andralojc, M. R. Hanson and M. A. Parry. 2016. "Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2." *Plant J* 85(1): 148-160.

Ohkawa, H., H. B. Pakrasi and T. Ogawa. 2000. "Two types of functionally distinct NAD(P)H dehydrogenases in *Synechocystis* sp. strain PCC 6803." *J Biol Chem* 275(41): 31630-31634.

Omata, T., G. D. Price, M. R. Badger, M. Okamura, S. Gohta and T. Ogawa. 1999. "Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942." *Proc Natl Acad Sci U S A* 96(23): 13571-13576.

Omata, T., Y. Takahashi, O. Yamaguchi and T. Nishimura. 2002. "Structure, function and regulation of the cyanobacterial high-affinity bicarbonate transporter, BCT1." *Functional plant biology* 29(3): 151-159.

Pang, A., M. J. Warren and R. W. Pickersgill. 2011. "Structure of PduT, a trimeric bacterial microcompartment protein with a 4Fe-4S cluster-binding site." *Acta*

Crystallogr D Biol Crystallogr 67(Pt 2): 91-96.

Pang, A., M. Liang, M. B. Prentice and R. W. Pickersgill. 2012. "Substrate channels revealed in the trimeric *Lactobacillus reuteri* bacterial microcompartment shell protein PduB." *Acta Crystallogr D Biol Crystallogr* 68(Pt 12): 1642-1652.

Pang, A., S. Frank, I. Brown, M. J. Warren and R. W. Pickersgill. 2014. "Structural insights into higher order assembly and function of the bacterial microcompartment protein PduA." *Journal of Biological Chemistry* 289(32): 22377-22384.

Parry, M. A., P. J. Andralojc, J. C. Scales, M. E. Salvucci, A. E. Carmo-Silva, H. Alonso and S. M. Whitney. 2012. "Rubisco activity and regulation as targets for crop improvement." *Journal of Experimental Botany* 64(3): 717-730.

Parsons, J. B., S. D. Dinesh, E. Deery, H. K. Leech, A. A. Brindley, D. Heldt, S. Frank, C. M. Smales, H. Lünsdorf and A. Rambach. 2008. "Biochemical and structural insights into bacterial organelle form and biogenesis." *Journal of biological chemistry* 283(21): 14366-14375.

Parsons, J. B., S. Frank, D. Bhella, M. Liang, M. B. Prentice, D. P. Mulvihill and M. J. Warren 2010. "Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement." *Mol Cell* 38(2): 305-315.

Pengelly, J. J., B. Forster, S. von Caemmerer, M. R. Badger, G. D. Price and S. M. Whitney 2014. "Transplastomic integration of a cyanobacterial bicarbonate transporter into tobacco chloroplasts." *J Exp Bot* 65(12): 3071-3080.

Pitts, A. C., L. R. Tuck, A. Faulds-Pain, R. J. Lewis and J. Marles-Wright 2012. "Structural insight into the *Clostridium difficile* ethanolamine utilisation microcompartment." *PloS one* 7(10): e48360.

Price, G. D. 2011a. "Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism." *Photosynth Res* 109(1-3): 47-57.

Price, G. D. and M. R. Badger. 1989. "Expression of Human Carbonic Anhydrase in the Cyanobacterium *Synechococcus* PCC7942 Creates a High CO_2-Requiring Phenotype: Evidence for a Central Role for Carboxysomes in the CO_2 Concentrating Mechanism." *Plant Physiol* 91(2): 505-513.

Price, G. D. and S. M. Howitt. 2011b. "The cyanobacterial bicarbonate transporter BicA: its physiological role and the implications of structural similarities with human SLC26 transporters." *Biochem Cell Biol* 89(2): 178-188.

Price, G. D., F. J. Woodger, M. R. Badger, S. M. Howitt and L. Tucker. 2004. "Identification of a SulP-type bicarbonate transporter in marine cyanobacteria." *Proc Natl Acad Sci U S A* 101(52): 18228-18233.

Price, G. D., J. J. Pengelly, B. Forster, J. Du, S. M. Whitney, S. von Caemmerer, M. R.

Badger, S. M. Howitt and J. R. Evans. 2012. "The cyanobacterial CCM as a source of genes for improving photosynthetic CO_2 fixation in crop species." *Journal of experimental botany* 64(3): 753-768.

Price, G. D., M. C. Shelden and S. M. Howitt. 2011c. "Membrane topology of the cyanobacterial bicarbonate transporter, SbtA, and identification of potential regulatory loops." *Mol Membr Biol* 28(5): 265-275.

Price, G. D., M. R. Badger and S. von Caemmerer. 2011d. "The prospect of using cyanobacterial bicarbonate transporters to improve leaf photosynthesis in C3 crop plants." *Plant Physiology* 155(1): 20-26.

Price, G. D., M. R. Badger, F. J. Woodger and B. M. Long. 2008. "Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants." *J Exp Bot* 59(7): 1441-1461.

Price, G. D., S.-i. Maeda, T. Omata and M. R. Badger. 2002. "Modes of active inorganic carbon uptake in the cyanobacterium, *Synechococcus* sp. PCC7942." *Functional Plant Biology* 29(3): 131-149.

Rae, B. D., B. Forster, M. R. Badger and G. D. Price. 2011. "The CO_2-concentrating mechanism of *Synechococcus* WH5701 is composed of native and horizontally-acquired components." *Photosynth Res* 109(1-3): 59-72.

Rae, B. D., B. M. Long, M. R. Badger and G. D. Price. 2013. "Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria." *Microbiol Mol Biol Rev* 77(3): 357-379.

Roberts, E. W., F. Cai, C. A. Kerfeld, G. C. Cannon and S. Heinhorst. 2012. "Isolation and characterization of the *Prochlorococcus* carboxysome reveal the presence of the novel shell protein CsoS1D." *J Bacteriol* 194(4): 787-795.

Sage, R. F. and X. G. Zhu. 2011. "Exploiting the engine of C_4 photosynthesis." *J Exp Bot* 62(9): 2989-3000.

Sagermann, M., A. Ohtaki and K. Nikolakakis. 2009. "Crystal structure of the EutL shell protein of the ethanolamine ammonia lyase microcompartment." *Proc Natl Acad Sci U S A* 106(22): 8883-8887.

Samal, A. and O. C. Martin. 2013. "Shining fresh light on the evolution of photosynthesis." *Elife* 2: e01403.

Samborska, B. and M. S. Kimber. 2012. "A dodecameric CcmK2 structure suggests beta-carboxysomal shell facets have a double-layered organization." *Structure* 20(8): 1353-1362.

Savage, D. F., B. Afonso, A. H. Chen and P. A. Silver. 2010. "Spatially ordered dynamics of the bacterial carbon fixation machinery." *Science* 327(5970):

1258-1261.

Schneider, G., Y. Lindqvist and C. I. Branden. 1992. "Rubisco: structure and mechanism." *Annu Rev Biophys Biomol Struct* 21: 119-143.

Shelden, M. C., S. M. Howitt and G. D. Price. 2010. "Membrane topology of the cyanobacterial bicarbonate transporter, BicA, a member of the SulP (SLC26A) family." *Molecular membrane biology* 27(1): 12-22.

Shibata, M., H. Katoh, M. Sonoda, H. Ohkawa, M. Shimoyama, H. Fukuzawa, A. Kaplan and T. Ogawa. 2002. "Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: function and phylogenetic analysis." *J Biol Chem* 277(21): 18658-18664.

Shih, P. M., A. Occhialini, J. C. Cameron, P. J. Andralojc, M. A. Parry and C. A. Kerfeld. 2016. "Biochemical characterization of predicted Precambrian Rubisco." *Nature communications* 7: 10382.

Shively, J. M., F. Ball, D. H. Brown and R. E. Saunders. 1973. "Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*." *Science* 182(4112): 584-586.

Sinha, S., S. Cheng, Y. W. Sung, D. E. McNamara, M. R. Sawaya, T. O. Yeates and T. A. Bobik. 2014. "Alanine scanning mutagenesis identifies an asparagine-arginine-lysine triad essential to assembly of the shell of the Pdu microcompartment." *J Mol Biol* 426(12): 2328-2345.

Sun, Y., S. Casella, Y. Fang, F. Huang, M. Faulkner, S. Barrett and L. N. Liu. 2016. "Light Modulates the Biosynthesis and Organization of Cyanobacterial Carbon Fixation Machinery through Photosynthetic Electron Flow." *Plant Physiol* 171(1): 530-541.

Sutter, M., B. Greber, C. Aussignargues and C. A. Kerfeld. 2017. "Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell." *Science* 356(6344): 1293-1297.

Sutter, M., M. Faulkner, C. Aussignargues, B. C. Paasch, S. Barrett, C. A. Kerfeld and L. N. Liu. 2016. "Visualization of Bacterial Microcompartment Facet Assembly Using High-Speed Atomic Force Microscopy." *Nano Lett* 16(3): 1590-1595.

Sutter, M., S. C. Wilson, S. Deutsch and C. A. Kerfeld. 2013. "Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species." *Photosynth Res* 118(1-2): 9-16.

Tabita, F. R., T. E. Hanson, H. Li, S. Satagopan, J. Singh and S. Chan. 2007. "Function, structure, and evolution of the Rubisco-like proteins and their Rubisco homologs." *Microbiol Mol Biol Rev* 71(4): 576-599.

Takenoya, M., K. Nikolakakis and M. Sagermann. 2010. "Crystallographic insights into

the pore structures and mechanisms of the EutL and EutM shell proteins of the ethanolamine-utilizing microcompartment of *Escherichia coli*." *Journal of bacteriology* 192(22): 6056-6063.

Tanaka, S., C. A. Kerfeld, M. R. Sawaya, F. Cai, S. Heinhorst, G. C. Cannon and T. O. Yeates. 2008. "Atomic-level models of the bacterial carboxysome shell." *Science* 319(5866): 1083-1086.

Tanaka, S., M. R. Sawaya and T. O. Yeates. 2010. "Structure and mechanisms of a protein-based organelle in *Escherichia coli*." *Science* 327(5961): 81-84.

Tanaka, S., M. R. Sawaya, M. Phillips and T. O. Yeates. 2009. "Insights from multiple structures of the shell proteins from the beta-carboxysome." *Protein Sci* 18(1): 108-120.

Taylor T. C., I. Andersson. 1997. The structure of the complex between Rubisco and its natural substrate ribulose 1,5-bisphosphate. *J. Mol. Biol*. 265, 432–444.

Tcherkez, G. G., G. D. Farquhar and T. J. Andrews. 2006. "Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized." *Proc Natl Acad Sci U S A* 103(19): 7246-7251.

Thompson, M. C. and T. O. Yeates. 2014. "A challenging interpretation of a hexagonally layered protein structure." *Acta Crystallogr D Biol Crystallogr* 70(Pt1): 203-208.

Tsai, Y. C., M. C. Lapina, S. Bhushan and O. Mueller-Cajar. 2015. "Identification and characterization of multiple Rubisco activases in chemoautotrophic bacteria." *Nat Commun* 6: 8883.

Tsai, Y., M. R. Sawaya, G. C. Cannon, F. Cai, E. B. Williams, S. Heinhorst, C. A. Kerfeld and T. O. Yeates. 2007. "Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome." *PLoS Biol* 5(6): e144.

Volokita, M., D. Zenvirth, A. Kaplan and L. Reinhold. 1984. "Nature of the Inorganic Carbon Species Actively Taken Up by the Cyanobacterium *Anabaena variabilis*." *Plant Physiol* 76(3): 599-602.

von Caemmerer, S. and J. R. Evans. 2010. "Enhancing C3 photosynthesis." *Plant Physiol* 154(2): 589-592.

von Caemmerer, S., W. P. Quick and R. T. Furbank. 2012. "The development of C_4 rice: current progress and future challenges." *Science* 336(6089): 1671-1672.

Wang, C., L. Guo, Y. Li and Z. Wang. 2012. "Systematic comparison of C_3 and C_4 plants based on metabolic network analysis." *BMC systems biology* Vol. 6. No. 2 BioMed Central.

Watson, G. M. and F. R. Tabita. 1997. "Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigation." *FEMS Microbiol Lett* 146(1): 13-22.

Whitehead, L., B. M. Long, G. D. Price and M. R. Badger. 2014. "Comparing the in vivo function of alpha-carboxysomes and beta-carboxysomes in two model cyanobacteria." *Plant Physiol* 165(1): 398-411.

Whitney, S. M., R. L. Houtz and H. Alonso. 2011 "Advancing our understanding and capacity to engineer nature's CO_2-sequestering enzyme, Rubisco." *Plant Physiology* 155(1): 27-35.

Xu, M., T. Ogawa, H. B. Pakrasi and H. Mi. 2008. "Identification and localization of the CupB protein involved in constitutive CO_2 uptake in the cyanobacterium, *Synechocystis sp.* strain PCC 6803." *Plant Cell Physiol* 49(6): 994-997.

Yeates, T. O., C. A. Kerfeld, S. Heinhorst, G. C. Cannon and J. M. Shively. 2008. "Protein-based organelles in bacteria: carboxysomes and related microcompartments." *Nat Rev Microbiol* 6(9): 681-691.

Yeates, T. O., C. S. Crowley and S. Tanaka. 2010. "Bacterial microcompartment organelles: protein shell structure and evolution." *Annu Rev Biophys* 39: 185-205.

Yeates, T. O., J. Jorda and T. A. Bobik. 2013. "The shells of BMC-type microcompartment organelles in bacteria." *J Mol Microbiol Biotechnol* 23(4-5): 290-299.

Yeates, T. O., M. C. Thompson and T. A. Bobik. 2011. "The protein shells of bacterial microcompartment organelles." *Curr Opin Struct Biol* 21(2): 223-231.

Zarzycki, J., S. D. Axen, J. N. Kinney and C. A. Kerfeld. 2012. "Cyanobacterial-based approaches to improving photosynthesis in plants." *Journal of experimental botany* 64(3): 787-798.

Zhang, P., N. Battchikova, T. Jansen, J. Appel, T. Ogawa and E. M. Aro. 2004. "Expression and functional roles of the two distinct NDH-1 complexes and the carbon acquisition complex NdhD3/NdhF3/CupA/Sll1735 in *Synechocystis sp* PCC 6803." *Plant Cell* 16(12): 3326-3340.

Chapter 2

Materials and Methods

2.1 Medium and Culture of *E. coli* and cyanobacteria

The *E. coli* TOP10 (One Shot™ TOP10 Chemically Competent *E. coli* Thermofisher Scientific C404010) or BL21(DE3) [One Shot™ BL21(DE3) Chemically Competent *E. coli* Thermofisher Scientific C600003] cells grown aerobically in lysogeny broth (LB) or LB agar plates (Sigma, UK). Medium supplements were used at the following final concentrations: 50 $\mu g \cdot mL^{-1}$ for kanamycin, spectinomycin and chloramphenicol, 100 $\mu g \cdot mL^{-1}$ ampicillin.

The cultures of *Synechococcus elongatus* strain PCC 7942 (Koksharova et al., 2006) were grown at 30ºC in BG11 medium (Table 2-1) in culture flasks with constant shaking or on BG11 1.5% (weight/volume) agar plates with white light intensity about 50 $\mu E \cdot m^{-2} \cdot s^{-1}$ (Sun et al., 2016).

Table 2-1. Stock solutions of BG11 (Castenholz 1988)

100 x BG11:	g/L
$NaNO_3$	149.6
$MgSO_4 \cdot 7H_2O$	7.49
$CaCl_2$	3.6
Citric acid	0.6
Na_2EDTA	1.12 mL 0.25M solution, pH 8.0
Trace elements:	g/100 mL
H_3BO_3	0.286
$MnCl_2 \cdot 4H_2O$	0.181
$ZnSO_4 \cdot 7H_2O$	0.022
$Na_2MoO_4 \cdot 2H_2O$	0.039
$CuSO_4 \cdot 5H_2O$	0.008
$Co(NO_3)_2 \cdot 6H_2O$	0.005
Iron stock:	g/100 mL
Ferric citrate	0.6
Phosphate stock	g/100 mL
K_2HPO_4	3.05
Na_2CO_3 stock	g/100 mL
Na_2CO_3	2.0
100 x TES buffer	
TES (1M)	22.9g, NaOH to pH 8.2

For 1 litre BG11 liquid medium, add 10 mL 100 x BG11, 1 mL trace elements and 1 mL iron stock, autoclave and cool to 50ºC, then add 1 mL phosphate stock and 1 mL Na_2CO_3 stock (sterilized by 0.22 μm filter separately), add antibiotics if required.

For 1 litre BG11 agar plates, add 10 mL 100 x BG11, 1 mL trace elements, 1 mL iron stock, 10 mL 100 x TES buffer, 3 g sodium thiosulphate and 15 g agar, autoclave and cool to 50ºC, add 1 mL phosphate stock and 1 mL Na_2CO_3 stock (sterilized by 0.22 μm filter separately), add antibiotics if required, mix and pour plates.

2.2 Genomic extraction, plasmid minipreparation and competent cell preparation

The genomic DNA of *Synechococcus elongatus* PCC 7942*, Halothiobacillus neapolitanus* (Cannon et al., 2003; Hutchinson et al., 1965)*,* and *Salmonella Typhimurium* LT2 (McClelland et al, 2001) was extracted from cell culture by GenElute™ Bacterial Genomic DNA Kit (Sigma) (Vázquez-Laslop et. al., 2006). All the plasmids were mini prepared by the GeneJET Plasmid Miniprep Kit (Thermo-Fisher, UK) (Birnboim et al., 1979; Vogelstein et al., 1979).

For the chemically competent cells, 250 mL of OD_{600}~0.4 *E.coli* cells were placed on ice for 20 min, and the cells harvested by centrifugation at 3000 g for 15 min at 4ºC. The supernatant was decanted and the pellet was ently resuspended in about 20 mL of ice-cold 100 mM $MgCl_2$. The bacteria were pelleted at 3000 g at 4°C for 10 min, resuspended in 15 mL of 100 mM $CaCl_2$, and incubated on ice for 30 min. After

spinning down again, the supernatant was discarded, and the pellets gently resuspended in 4 mL of 0.1 M $CaCl_2$ with 20% (v/v) glycerol solution. The cells were stored in 50 µl aliquots and frozen with liquid nitrogen at −80°C (Sambrook et al., 1989).

2.3 Preparation of Plasmids

In Chapter 3, genes were cloned from the genomic DNA of *Synechococcus elongatus* PCC 7942 by CloneAmp™ HiFi PCR Premix (Clontech), synthesized genes were brought from Sigma-Aldrich (Holliday et al., 2014). Primers were designed to contain In-Fusion (Zhu et al., 2007) recombination sites appropriate for the assembly of the desired construct (Table 2-2). Every operon contains 50 bp upstream and 20 bp downstream to preserve the native promoter and ribosome-binding site sequence of all the genes. The operons were inserted into the pETM11 vector digested at *EcoR* I and *Xho* I sites or the pAM2991 vector (a gift from Susan Golden from Texas A&M University, Addgene plasmid # 40248) (Ivleva et al., 2005) at *EcoR* I and *BamH* I sites by In-Fusion® HD Cloning (Clontech) (Zhu et al., 2007). Genes encoding RbcL-eGFP-RbcS were cloned from the genomic DNA of the RbcL-eGFP (enhanced Green Fluorescent Protein) Syn7942 mutant (Sun et al., 2016) and inserted into the pTTQ18 vector (Addgene plasmid # 69122) (Stark et al., 1987) at *EcoR* I and *Xba* I sites; the same strategy was used for cloning and inserting CcmK4-GFP at *EcoR* I and *Kpn* I sites. Genes encoding CsoS1A and CbbL were cloned from the genomic

DNA of *H. neapolitanus,* fused with the genes of CFP (Cyan Fluorescent Protein) and eYFP (enhanced Yellow Fluorescent Protein), respectively, and inserted into pAM2991 by In-Fusion cloning. The *pduA* gene was cloned from the genome of *S. Typhimurium* LT2, fused with superfolder GFP (sfGFP), and inserted into pAM2991 by In-Fusion cloning.

In Chapter 4, CsoS1D plasmids were constructed using In-Fusion® HD Cloning Plus (Zhu et al., 2007) and CloneAmp™ HiFi PCR Premix. The operons for *H. neapolitanus cso* and *csoS1D* was amplified from the pHnCBS1D plasmid (Addgene plasmid # 52065) (Bonacci et al., 2012) containing nine genes of the *cso* operon and *csoS1D*. The pAM2991 plasmid for CsoS1D plasmid was digested with *BamH* I and *EcoR* I. The CsoS1DCbbQ plasmid was obtained from Prof G Dean Price (Australian National University).

In Chapter 5, the CsoS1D plasmid from Chapter 4 was transformed into an electrocompetent BW25113 strain (Baba et al.,2006) which carries the λ Red expression plasmid pKD46 (Datsenko et al, 2000). The *aadA* gene (confers streptomycin resistance) was amplified from the pIJ778 plasmid (Gust et al., 2003) with 39bp overlap as the same as the DNA sequence in front of *cbbL* and after *cbbS*, and then electroporation transformed into the same BW25113 cell. The colonies were selected on the spectinomycin LB plate to get the Δ*cbbL*:*cbbS* plasmid. Then spectinomycin was cut off with utilising Flp-FRT recombination (Zhu et al., 1995) and the pCP20 plasmid (Cherepanov et al., 1995).

All the restriction endonucleases were purchased from New England Biolabs (UK). Other reagents were purchased from Sigma-Aldrich (US) or Thermo-Fisher (UK).

Table 2-2. Primer lists

Primers	Sequences 5'-3'	Description	Source
Long fragment F:	CCCCTGCCAGTCTCTTCTGATTAAGGTTTGCTTCTTGCTAGAGAC	To amplify the operons of β-carboxysome for pLFbC901-pETM11 plasmid, which is used in Chapter 3 to express full β-carboxysome in *E.coli*. In-fusion primers.	PCR template from Syn7942 genomic DNA and synthesized genes.
Long fragment R:	TCAACATATTCTACTATCAAGACAAATCAGGCTTTAG		
Short fragment F:	TTTGTCTTGATAGTAGAATATGTTGATGGCTCTTC		
Short fragment R:	ATCTCAGTGGTGGTGGTGGTGGTGCAATCAACGTGTTGAACAATTTC		
6k F	CCCCTGCCAGTCTCTTCTGATTAAGGTTTGCTTCTTGCTAGAGAC	To amplify the main operon of β-carboxysome for pLFbC601-pETM11 plasmid, which is used in Chapter 3 to express partial β-carboxysome in *E.coli*. In-fusion primers.	PCR template from Syn7942 genomic DNA.
6k R	ATCTCAGTGGTGGTGGTGGTGGTGCCTATCAAGACAAATCAGGCTTTAG		
RbcL F	ACAGGAAACAGCGATGCCCAAGACGCAATCTGCC	To amplify the fragment for *rbcL-eGFP-rbcS*-pTTQ18 plasmid used in Chapter 3. In-fusion primers.	PCR template from Syn7942 genomic DNA (Sun et al., 2016).
RbcL R	GCCTGCAGGTCGACTTTAGTAGCGGCCGGGACG		
ccmK4 F	GCGCGAATTCGATGTCTCAGCAGGCAATTGG	To amplify the fragment for *ccmK4-eGFP*-pTTQ18 plasmid used in Chapter 3. In-fusion primers.	PCR template from Syn7942 genomic DNA (Sun et al.,
ccmK4 R	GCGCGGTACCATTCCGGGGATCCGTCGACC		

			2016).
β-cb-pAM2991 F	AGGAAACAGACCATGGGTTTGCTTCTTGCTAGAG	To amplify the fragment for pLFbC901- pAM2991 plasmid which can express β-cb in Syn7942, used in Chapter 3. In-fusion primers.	PCR template from pLFbC901-pETM11.
β-cb-pAM2991 R	CAGGTCGACTCTAGAGAATCAACGTGTTGAACAATTTC		
CsoS1A F	AGGAAACAGACCATGGAGGCTTGTTCCGGTTCTTTG	To amplify the fragments for *csoS1A-CFP*- pAM2991 plasmid which can express α-carboxysome protein in Syn7942, used in Chapter 3. In-fusion primers.	PCR template from genomic DNA and CFP-pIJ778 plasmid (per-existing in this lab).
CsoS1A R:	CTCCGGGCCCGGCAGTTAGGCTTGTGGCGCCTT		
CFP F	CTGCCGGGCCCGGAGCTG		
CFP R	CAGGTCGACTCTAGAG TTAAGCGCTTACGTAGAGC		
CbbL F	AGGAAACAGACCATGGATGGCAGTTAAAAAGTATAGTG	To amplify the fragments for *cbbL-YFP*- pAM2991which can express α-carboxysome protein in Syn7942, used in Chapter 3. In-fusion primers.	PCR template from *H. neapolitanus* genomic DNA and YFP-pIJ778 (per-existing in this lab).
CbbL R	CTCCGGGCCCGGCAGTCAACGATTTTGAGTGTCGA		
YFP F	CTGCCGGGCCCGGAGCTG		
YFP R	CAGGTCGACTCTAGAGTTACTTGTACAGCTCGTC		
pduA F	ACACAGGAAACAGACCATGGCCCAACTATCGGAACACTC	To amplify the fragments for *pduA-sfGFP*- pAM2991 which can express PDU protein in Syn7942, used in Chapter 3. In-fusion primers.	PCR template from *S.Typhimurium* LT2 genomic DNA and sfGFP
pduA R	GGAGCCAGCGGATCCTTGGCTAATTCCCTTCGGTA		
sfGFP F	GCCAAGGATCCGCTG		

	GCTCCGCTGC		plasmid (per-existing in this lab).
sfGFP R	CCTGCAGGTCGACTCTAGAGAAGGAGGATATTCATATGTTATTTG		
CsoS1D-F	AGGAAACAGACCATGGATGGCAGTTAAAAAGTATAGTGC	To amplify the α-carboxysome whole operon for *csoS1D*-pAM2991 plasmid in Chapter 4, which can express α-carboxysome in *E.coli*.	PCR template from pHnCBS1D plasmid (Bonacci et al., 2012)
CsoS1D-R	CAGGTCGACTCTAGAGTTAGAACCCTTCAGCGCG		
CsoS1D-1	ACGGCGTACTGCTGCACATC	The DNA sequencing primers used in Chapter 4 to detect the sequence of CsoS1D and CsoS1DCbbQ plasmid.	
CsoS1D-2	AGCGTGTCGTAGTGCGTATC		
CsoS1D-3	TCCAGGACGCAATGGCTATC		
CsoS1D-4	GCGGTTCGTGTCGTTCAATC		
CsoS1D-5	CGGCTTGATTACAGGAACCC		
CsoS1D-6	TGGACATACGCTGCCCGAAG		
CsoS1D-7	CAACAACCTCTCGCAAATCG		
CsoS1D-8	GAGGGATGTTGGGTCTTTAC		
CsoS1D-9	GCGTGTCCACTCAGAAGTAG		
pAM2991 F	TCGGAAGCTGTGGTATGGCT		
pAM2991 R	AGATAAAACGAAAGGCCCAG		
CsoCbbQ-10	ACCTTGATCAATCGCCAGTT		
CsoCbbQ-11	TTAAAAGAACGTTTTG		

	ACGA		
Spec-F	AGCGGATAACAATTTCACACAGGAAACAGACCATGGATGATTCCGGGGATCCGTCGACC	To amplify *aadA* gene, used in Chapter 5. Recombination primers.	PCR template from pIJ778 plasmid.
Spec-R	GCGTAGAGAAACGCACAGCGCAATGACAGACTTGACTTATGTAGGCTGGAGCTGCTTC		
CbbL F	ATGGCAGTTAAAAAGTATAG	Primers can amplify CbbL-CbbS operon, used for detection in Chapter 5.	
CbbS R	TTAGTTGCCGCGGTAGACCA		
CsoS3 F	TCATGAACACCCGTAACAC	Primers can amplify CsoS3-CsoS4B operon, used for detection in Chapter 5.	
CsoS4B R	TAGCGATGGTCAAGTTAC		

2.4 Heterologous expression of α-, β-carboxysomes and α-shell

BL21(DE3) cells transformed with the expression plasmids were grown in LB media at 37ºC until OD_{600} reaches 0.3–0.5, then isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added into the cultures at a final concentration of 50 μM for protein introduction (Hansen et al., 1998) at 18ºC overnight.

Syn7942 wild-type and RbcL-eGFP and CcmK4-eGFP mutants (Sun et al., 2016) were transformed with the generated pAM2991 vectors described above. Cells from single colonies were grown in BG11 media at the constant light until the middle of the exponential growth phase and then induced proteins with 100 μM IPTG overnight at 30°C.

2.5 Purification of synthetic carboxysomes and shells

The BL21(DE3) competent cells were transformed with plasmids expressing α-, β-carboxysomes or α-carboxysome shells. 10 mL of cells were grown overnight at 37ºC. The cultures were then diluted 1:100 in 1 L flasks of LB and incubated at 37ºC with shaking until OD_{600} reaches 0.4, and then were induced with 50 μM IPTG for overnight at 18ºC. Cells were harvested by centrifugation at 5,000 g for 10 min at 4ºC.

For β-carboxysome isolation, cell pellets were washed once with TE buffer (10 mM Tris pH 8.0, 0.1 mM EDTA), and was then resuspended into 18 mL TE buffer and 2 mL CelLytic™ B Cell Lysis Reagent (Sigma-Aldrich, US) with 1% Protease Inhibitor Cocktails (Thermo-Fisher, UK), and then cell breakage was achieved using a Stansted pressure cell homogenizer three times (Stansted Fluid Power, UK). The cell lysate was then treated with 1% n-dodecyl β-maltoside (Sigma-Aldrich, UK) for 1 hour at 4ºC.

For α-carboxysome and α-carboxysome shell isolation, cell pellets were washed once with TEMB buffer (10 mM Tris-pH8.0, 10 mM $MgCl_2$, 1mM EDTA, and 20 mM $NaHCO_3$) and then resuspended in 18 mL buffer TEMB and 2 mL CelLytic™ B Cell Lysis Reagent (Sigma-Aldrich, UK) and 1% Protease Inhibitor Cocktails (Thermo-Fisher, UK), and then cell breakage was achieved using a Stansted pressure cell homogenizer three times (Stansted Fluid Power, UK).

After the cell breakage, cellular debris was removed by centrifuge at 12,000 g for 10

min. The supernatants were centrifuged at 50,000 g for 20 min to enrich synthetic carboxysomes or carboxysome shells. The generated pellet was gently resuspended in 2 mL TE or TEMB buffer and was centrifuged at 3,000 g 1min. For β-carboxysomes, the 2 mL supernatant was applied to a 9-mL sucrose gradient made of 1.5 mL of each 10%, 20%, 30%, 40%, 50% and 60% sucrose in TE buffer. For α-carboxysomes and α-carboxysome shells, the 2 mL supernatant was applied to a 10-mL sucrose gradient made of 2 mL of each 10%, 20%, 30%, 40% and 50% sucrose in TEMB buffer.

The gradient was centrifuged at 105,000g for 30 min at 4ºC. Fractions of 1~1.5 mL were collected and analyzed for SDS-PAGE. The carboxysome-enriched fractions were diluted 20 times using TE or TEMB buffers and were centrifuged at 50,000 g for 20 min to pellet the carboxysomes. The pellets were then resuspended in 0.5 mL TE or TEMB buffer and stored at 4ºC for short-term or nitrogen frozen and stored at -80ºC for long-term.

2.6 SDS-PAGE

BL21(DE3) soluble fractions were prepared by sonication at 4°C and then mixed with 4x SDS-PAGE sample loading buffer (250 mM Tris–pH 6.8, 8% (w/v) SDS, 0.2% (w/v) bromophenol blue, 40% (v/v) glycerol, 20% (v/v) β-mercaptoethanol). For protein samples, 20 mg of the isolated carboxysomes and carboxysomes shell proteins were mixed with 4x SDS-PAGE sample loading buffer. After 95ºC heating for 10 mins, samples were centrifuged at 12,000 g for 1 min and then loaded on 15% (v/v)

denaturing SDS-PAGE gels. 3 μL of unstained protein ladder (PageRuler™ Plus Prestained Protein Ladder, 10-250 kDa from Thermo-Fishe) was loaded as size standards. SDS-PAGE was performed at 170 V with Laemmli electrophoresis running buffer (25 mM Tris, 192 mM glycine, and 1% SDS). The gels were stained by Coomassie blue stain buffer (0.25% Coomassie Brilliant Blue R-250, 20% Methanol, 10% Acetic Acid) and distained by destaining buffer (20% Methanol, 10% Acetic Acid).

Table 2-3. Recipes for SDS-PAGE gel

15% (v/v) denaturing SDS-PAGE gel:	total volume 5 mL
Acrylamide/bis (30% 37.5:1; Severn biotech)	2.5 mL
Tris (1.5 M, pH 8.8)	1.25 mL
SDS, 10%	50 μL
Tetramethylethylenediamine (TEMED) (Thermo-Fisher)	50 μL
Ammonium persulfate (APS), 10%	5 μL
MQ H2O	1.15 mL
4% (v/v) stacking gel:	total volume 2.5 mL
Acrylamide/bis (30% 37.5:1; Severn biotech)	0.335 mL
Tris (0.5 M, pH 6.8)	0.625 mL
SDS, 10%	25 μL
Tetramethylethylenediamine (TEMED) (Thermo-Fisher)	25 μL
Ammonium persulfate (APS), 10%	2.5 μL
MQ H2O	1.49 mL

2.7 Immunoblot Analysis

The samples for western blot were loaded on 15% (v/v) denaturing SDS-PAGE gels with PageRuler^{TM} Plus Prestained Protein Ladder, 10 to 250 kDa from Thermo-Fisher. 4 pieces of 3 MM filter paper and a piece of PVDF membrane (Bio-Rad, US) were per-cut, and then wetted the PVDF membrane in 100% methanol for 30 sec, drained and equilibrated the membrane immediately with transference buffer (25 mM Tris base, 150 mM glycine, 10% methanol) and shaking for 5 minutes. The gel was also soaked in transference buffer for 15 sec and then assembled the blot 'sandwich'. The transfer was produced at 90 V for 45 min at 4ºC in transference buffer.

After transfer, PVDF membrane was rinsed with TBS buffer (50 mM Tris-pH7.5, 150 mM NaCl) for 5 min and proceeded to immunoblot. First, block the membrane with 5% milk in TBS (blocking buffer) for 1 h at room temperature or overnight at 4ºC, and wash the membrane with TBST buffer (50 mM Tris, 150 mM NaCl, 0.1% Tween 20) for 5 min. Then the membrane is incubated with primary antibody (rabbit polyclonal anti-RbcL dilute in 1:10,000 from Agrisera, Sweden or anti-CcmK2 antibody, a gift from Prof Kerfeld, Berkeley) in blocking buffer for 4 h at room temperature or overnight in cold room on a swing shaker. When the incubation was finished, the membrane washed in TBST buffer (3 times), 5 min each, and then incubate the membrane in secondary antibody (anti-rabbit immunoglobulin G dilute in 1:10,000 from GE Healthcare, US) in TBST buffer at room temperature for 1.5 h. Finally the membrane was washed in TBST buffer (3 times, 5 min each), followed by two washes with TBS buffer, 5 min each. Signals were visualized using the Clarity Western ECL

Substrate Kits (Bio-Rad, US) and images were taken by ImageQuant LAS 4000 (GE Healthcare Life Sciences).

2.8 Proteomic analysis

Whole cell proteomic analysis of β-carboxysomes was performed as follows: 20 mL of *E. coli* cells expressing synthetic carboxysome proteins were harvested in 25 mM NH_4HCO_3 buffer. After washing and sonication, the cell lysates in 25 mM NH_4HCO_3 bufferin in 1.5 mL tubes were subjected to mass spectrometry analysis. The isolated 50μg proteins of α-carboxysome and α-carboxysome shell from the sucrose fraction were washed with PBS buffer and followed the protocol as described previously (Faulkner et al., 2017).

A final concentration of 0.05% (w/v) of RapiGest SF (Waters Corporation,US) was added to the sample for 10 min incubation at 80ºC. The sample was then reduced with 3 mM dithiothreitol for 10 min at 60ºC, then 9 mM iodoacetamide was added for 30 min in the dark at room temperature, and sample digested with trypsin at 37ºC overnight. The digest was terminated by the adding of 1 μL trifluoroacetic acid (TFA). Data-dependent LC-MS/MS analysis was carried on the QExactive quadrupole-Orbitrap mass spectrometer with the Dionex Ultimate 3000 RSLC nano-liquid chromatograph (Hemel Hempstead, UK). 2 μL of each sample was loaded onto the trapping column (Acclaim PepMap 100 C18, 75 μm × 2 cm, 3 μm packing material, 100 Å) in 0.1% TFA, 2% acetonitrile H_2O, and set in line with the analytical

column (EASY-Spray PepMap RSLC C18, 75 µm × 50 cm, 2 µm packing material, 100 Å). Peptides were eluted by a linear gradient of 96.2% buffer A (0.1% formic acid): 3.8% buffer B (0.1% formic acid in water: acetonitrile 80: 20, v/v) to 50% buffer A: 50% buffer B over 30 min at 300 nL min^{-1}. The mass spectrometry analysis was carried in data-dependentacquisition (DDA) mode with survey scans between m/z 300–2000 acquired at a mass resolution of 70 000 (FWHM) at m/z 200. The maximum injection time was 250 ms and the automatic gain control as $1e^6$. The peptides were fragmented by higher-energy collisional dissociation using a normalized collision energy of 30%. Dynamic exclusion of m/z values to prevent repeated fragmentation of the same peptide was used with an exclusion time of 20 sec.

The raw data file was analyzed by Progenesis QI for Proteomics (Version 3.0 Nonlinear Dynamics, Newcastle upon Tyne UK, a Waters Company). Peak picking parameters were applied with sensitivity set to maximum and features with charges of 2+ to 7+ were retained. A Mascot Generic File, created by Progenesis, was searched against the Syn7942 database from UniProt or *H. neapolitanus* database from UniProt with the sequence of yeast enolase (UniProt: P00924) added. Trypsin was specified as the protease with one missed cleavage allowed and with fixed carbamidomethyl modification for cysteine and variable oxidation modification for methionine. A precursor mass tolerance of 10 ppm and a fragment ion mass tolerance of 0.01 Da were applied. The results were then filtered to obtain a peptide false discovery rate of 1%. Protein quantification was calculated using Hi3 methodology using yeast enolase (50 fmol μL^{-1}) as a standard protein.

2.9 Rubisco assay

For *in vivo* Rubisco assays, 120 µL of *E. coli* cells (OD_{600} = 4.0) or 120 µL of Syn7942 cells (OD_{750} = 4.0) in Rubisco assay buffer (100 mM EPPS, pH 8.0; 20 mM $MgCl_2$) were added into a 250 µL reaction. Radiometric assays for ^{14}C were performed following the previously described protocol (Price and Badger, 1989) with an additional cell permeabilization treatment (Schwarz et al., 1995). Rubisco activities of isolated synthetic β-carboxysome structures were determined as previously described (Faulkner et al., 2017). 5 µL of 1 $mg \cdot mL^{-1}$ isolated synthetic β-carboxysomes in Rubisco assay buffer was added. Samples were added into scintillation vials containing $NaH^{14}CO_3$ at a final concentration of 25 mM and incubated at 30°C for 2 min before the addition of D-ribulose 1,5-bisphosphate sodium salt hydrate (RuBP, Sigma Aldrich, US) at a final concentration of 5 mM. The reaction proceeded for 5 min and was terminated by adding 2:1 volume 10% formic acid. The samples were dried for at least 30 min at 95°C to remove unfixed ^{14}C, and then the fixed ^{14}C pellets were resuspend in ultra-pure water and 2 mL of scintillation cocktail (Ultima Gold XR, PerkinElmer, US). Radioactivity measurements were then captured by the scintillation counter (Tri-Carb, PerkinElmer, US). Raw readings were used to calculate the amount of fixed ^{14}C, aligned by control samples without providing RuBP (Yaqi, 2016; Faulkner et al., 2017) and then converted to the total carbon fixation rates. Results are presented as mean ± standard deviation (SD).

2.10 Fluorescence microscopy

The *E. coli* BL21(DE3) cells possessing RbcL-eGFP-RbcS-pTTQ18 and CcmK4-pTTQ18 with or without pLFbC901 were cultured in 37ºC until OD_{600} reaches 0.6 and then they were induced with 50 μM IPTG overnight at 18°C. Preparation of *E. coli* and Syn7942 cells for confocal microscopy were set by drying a drop of cell culture onto LB or BG-11 agar plates. Blocks of agar with the cells dried onto the surface were covered with a coverslip and placed under the microscope. The *E. coli* and Syn7942 cells were imaged using a Zeiss LSM710 or LSM780 with a 63× or 100× oil-immersion objective and excitation at 488 nm. Live-cell images were recorded from at least five biological repeat cultures. All images were captured with all pixels below saturation. Image analysis was carried out using ImageJ software (NIH Image). Automated analysis of the number of carboxysomes in Syn7942 cells was programmed into the image analysis software Image SXM (http://www.imagesxm.org.uk/), as described earlier (Sun et al., 2016).

2.11 Transmission electron microscopy

E. coli cells expressing β-, α-carboxysomes and α-carboxysome shell proteins were pelleted and fixed for 1 h with 2% paraformaldehyde and 2% glutaraldehyde in 0.1 M sodium cacodylate buffer at pH 7.2. Samples were stained for 1 hour with 2% osmium tetroxide and 1.5% potassium ferrocyanide, followed by staining with 1% thiocarbohydrazide for 20 min and 2% osmium tetroxide for 30 min. Afterward,

samples were washed with ultrapure water, stained with 1% uranyl acetate for overnight at 4°C. The samples were dehydrated with a series of increasing alcohol concentrations (30 to 100%) and followed with embed in resin. Thin sections of 70 nm were cut with a diamond knife.

The structures of isolated β-, α-carboxysomes and α-carboxysome shells for negative staining were prepared on glow-discharged, collodion-carbon-coated copper grids (200 mesh/inch). A droplet of the sample solution was placed on the grid for 5 min, excess liquid was then blotted off and the samples were fixed with 2% glutaraldehyde for 90 sec, followed by successive washes on three droplets of double-distilled water. Samples on the grids were stained with 3% uranyl acetate. Images were recorded using an FEI Tecnai G2 Spirit BioTWIN FEI transmission electron microscope. Image analysis was carried out using ImageJ software (NIH Image) and purified β-carboxysome were processed by Plot Profile analysis (Figure 3-4D).

2.12 Atomic force microscopy

Isolated β-, α-carboxysome and α-carboxysome shell samples were adsorbed onto freshly cleaved mica surface in TN buffer (10 mM Tris, 5 mM $NiCl_2$, pH 8.0) for 1 hour at room temperature, and then washed and imaged with TN buffer. AFM imaging was captured on a MultiMode 8 AFM with NanoScope V controller (Bruker, Santa Barbara, US) in peak force tapping mode in liquid at room temperature. The spring constant of 0.4 N m^{-1} AFM tips (Scanassyst air HR, Bruker, Santa Barbara, US) were used for

high-resolution imaging, and the tip spring constant was routinely calibrated. The average imaging force was 100 pN.

AFM experiments on β-, α-carboxysomes and α-carboxysome shell were carried out in solution to ensure their structural and functional integrity (Faulkner et al., 2017). Confocal-AFM images were operated using a NanoWizard 3 AFM (JPK) integrated with a Zeiss LSM880 confocal microscope. Samples were adsorbed on glass slides in adsorption buffer (10 mM Tris-HCl, 150 mM KCl, 25 mM $MgCl_2$, pH 7.5) for 10 mins at room temperature, and then washed with imaging buffer (10 mM Tris-HCl, 150 mM KCl, pH 7.5). Confocal images were captured using a 40× objective with 488 nm excitation. Particles with high-intensity GFP signal were imaged by AFM in Quantitative Imaging (QI) mode. The scanning force is ~100 pN. Image analysis was performed using JPK SPM Data Processing (JPK).

2.13 References

Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. 2006. "Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection". *Molecular Systems Biology*. 2: 2006.0008.

Birnboim H.C., and Doly, J. 1979. “A rapid alkaline lysis procedure for screening recombinant plasmid DNA.” *Nucleic Acids Res*. 7, 1513 -1522.

Bonacci, W., P. K. Teng, B. Afonso, H. Niederholtmeyer, P. Grob, P. A. Silver and D. F. Savage. 2012. "Modularity of a carbon-fixing protein organelle." *Proc Natl Acad Sci U S A* 109(2): 478-483.

Cannon, G. C., S. H. Baker, F. Soyer, D. R. Johnson, C. E. Bradburne, J. L. Mehlman, P. S. Davies, Q. L. Jiang, S. Heinhorst and J. M. Shively .2003.. "Organization of carboxysome genes in the thiobacilli." Curr Microbiol 46(2): 115-119.

Castenholz, R. W. 1988. [3] Culturing methods for cyanobacteria. *Methods in*

enzymology, 167: 68-93.

Cherepanov, P. P. and W. Wackernagel. 1995. "Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant." *Gene* 158(1): 9-14.

Datsenko, K. A. and B. L. Wanner. 2000. "One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products." *Proceedings of the National Academy of Sciences* 97(12): 6640-6645.

Faulkner, M., J. Rodriguez-Ramos, G. F. Dykes, S. V. Owen, S. Casella, D. M. Simpson, R. J. Beynon and L. N. Liu. 2017. "Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes." *Nanoscale* 9(30): 10662-10673.

Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. 2003. “PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin.” *Proc. Natl. Acad. Sci. USA* 100(4), 1541–1546.

Hansen, L. H., S. Knudsen and S. J. Sorensen. 1998. "The effect of the lacY gene on the induction of IPTG inducible promoters, studied in Escherichia coli and Pseudomonas fluorescens." *Curr Microbiol* 36(6): 341-347.

Holliday, M. J., F. Zhang, N. G. Isern, G. S. Armstrong and E. Z. Eisenmesser. 2014. "1H, 13C, and 15N backbone and side chain resonance assignments of thermophilic Geobacillus kaustophilus cyclophilin-A." *Biomol NMR Assign* 8(1): 23-27.

Hutchinson, M., K. I. Johnstone and D. White. 1965. "The taxonomy of certain thiobacilli." *J Gen Microbiol* 41(3): 357-366.

Ivleva, N. B., M. R. Bramlett, P. A. Lindahl and S. S. Golden. 2005. "LdpA: a component of the circadian clock senses redox state of the cell." *EMBO J* 24(6): 1202-1210.

Koksharova, O. A., I. Klint and U. Rasmussen. 2006. "The first protein map of Synechococcus sp. strain PCC 7942." *Mikrobiologiia* 75(6): 765-774.

McClelland, M., K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston and R. K. Wilson. 2001. "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2." *Nature* 413(6858): 852-856.

Price, G. and M. Badger. 1989. "Isolation and characterization of high CO_2-requiring-mutants of the cyanobacterium *Synechococcus* PCC7942: two phenotypes that accumulate inorganic carbon but are apparently unable to generate CO_2 within the carboxysome." *Plant Physiology* 91(2): 514-525.

Sambrook J. Fritsch E.F. and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Second edition. *Cold Spring Harbor Laboratory Press*. p.1.74.

Schwarz, R., L. Reinhold and A. Kaplan. 1995. "Low Activation State of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase in Carboxysome-Defective *Synechococcus* Mutants." *Plant Physiol* 108(1): 183-190.

Stark M. J.1987. “Multicopy expression vectors carrying the lac repressor gene for regulated high-level expression of genes in Escherichia coli.” *Gene.* 1987;51(2-3):255-67.

Sun, Y., S. Casella, Y. Fang, F. Huang, M. Faulkner, S. Barrett and L. N. Liu. 2016. "Light Modulates the Biosynthesis and Organization of Cyanobacterial Carbon Fixation Machinery through Photosynthetic Electron Flow." *Plant Physiol* 171(1): 530-541.

Vazquez-Laslop, N., H. Lee and A. A. Neyfakh (2006). "Increased persistence in Escherichia coli caused by controlled expression of toxins or other unrelated proteins." *J Bacteriol* 188(10): 3494-3497.

Vogelstein, B. and Gillespie, D. 1979. “Preparative and analytical purification of DNA from agarose.” *Proc. Natl. Acad. Sci. USA* 76, 615-619.

Zhu, B., G. Cai, E. O. Hall and G. J. Freeman (2007). "In-fusion assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations." *Biotechniques* 43(3): 354-359.

Zhu, X. D. and P. D. Sadowski. 1995. "Cleavage-dependent ligation by the FLP recombinase. Characterization of a mutant FLP protein with an alteration in a catalytic amino acid." *J Biol Chem* 270(39): 23044-23054.

Chapter 3

Engineering and modulating functional β-cyanobacterial CO_2-fixing organelles

3.1 Introduction

Subcellular compartmentalization is a hallmark of eukaryotic cells, allowing cells to perform and confine various chemical reactions in space and time, and provide a means for eliminating metabolic crosstalk and enhancing the efficiency of compartmentalized metabolic pathways (Giessen and Silver, 2016; Hammer and Avalos, 2017). Compartmentalization also occurs in the prokaryotic cytoplasm and membranes, such as photosynthetic membranes (Liu, 2016). A particular example is bacterial microcompartments (BMCs), which are distributed in at least 23 different bacterial phyla, sequestering enzymes within the specialized niches from the cytoplasm (Axen et al., 2014; Chowdhury et al., 2014; Kerfeld and Erbilgin, 2015). All BMCs are made entirely of proteins, including interior enzymes that catalyze sequential reactions and an encapsulating shell. The architecture of the BMC shell resembles an icosahedral viral capsid, with facets composed of hexameric and trimeric proteins and vertices capped by pentameric proteins (Kerfeld et al., 2005; Tanaka et al., 2008; Klein et al., 2009; Sutter et al., 2017). These multiple protein paralogs have specific permeability for the passage of metabolites. Three types of BMCs have been extensively studied: the carboxysomes for CO_2 fixation, the PDU BMCs for 1, 2-propanediol utilisation, and the EUT BMCs for ethanolamine utilisation (Bobik et al., 2015; Kerfeld and Erbilgin, 2015).

The first BMCs identified were the carboxysomes, the central machinery for CO_2 fixation in cyanobacteria (Figure 3-1A)(Shively et al., 1973). Composed of thousands

of protein subunits, carboxysomes encapsulate the CO_2-fixation enzyme Rubisco and carbonic anhydrase in a protein-based shell. Compartmentalization of these enzymes and the selective permeability of the shell that allows for the diffusion of HCO_3^- and diminishes CO_2 leakage ensure a high CO_2 concentration near Rubisco (Cai et al., 2009), thereby favouring the higher carboxylation rates of Rubisco within carboxysomes and enhanced carbon fixation. According to the Rubisco phylogeny, protein composition and assembly, carboxysomes can be classified as α-carboxysomes (possessing Form 1A Rubisco) and β-carboxysomes (containing plant-like Form 1B Rubisco) (Rae et al., 2013). It has been deduced that α-carboxysomes and β-carboxysomes embrace distinct assembly pathways. The α-carboxysome shell assembles concomitantly with aggregation of Rubisco (Iancu et al., 2010) or even without Rubisco packing (Baker et al., 1998; Menon et al., 2008), whereas *de novo* assembly of β-carboxysomes follows the “inside out” mode: the β-carboxysome shell begins assembly after RuBisCO aggregation (Chen et al., 2012; Cameron et al., 2013). The CO_2-fixing organelles in the model freshwater cyanobacterium *Synechococcus elongatus* PCC 7942 (Syn7942) are β-carboxysomes, which have specific spatial location and regulation in cyanobacterial cells (Savage et al., 2010; Sun et al., 2016). There are about 4~5 β-carboxysome per cell in Syn7942 under moderate light (ML, 50 $\mu E \cdot m^{-2} \cdot s^{-1}$) (Figure 3-1A), and the light and the roles of photosynthetic electron flow can control the numbers and positioning of β-carboxysome (Sun et al., 2016).

The β-carboxysome shell in Syn7942 is composed of the hexameric proteins major

shell protein CcmK2, and minor shell proteins CcmK3 and CcmK4 constructing shell facets (Kerfeld et al., 2005), CcmO as trimer protein (Rae et al., 2012) and the CcmL pentamers located at the vertices (Figure 3-1B)(Tanaka et al., 2008). Rubisco enzymes form a densely-packed paracrystalline array in the β-carboxysome lumen (Faulkner et al., 2017). The packing of Rubisco is mediated by the 35 kDa truncated version of CcmM (CcmM35)(Long et al., 2010; Long et al., 2011). The longer form of CcmM, Ccm58, contains an N-terminal domain with homology to γ-carbonic anhydrase and a C-terminal domain comprising three small subunit-like domains (SSUs) with homology to RbcS, the small subunit of Rubisco (Peña et al., 2010). CcmM interacts with the carboxysomal β-carbonic anhydrase (CcaA), Rubisco, and CcmN that acts as a bridge between CcmM and the shell (Figure 1-4A)(Long et al., 2007; Long et al., 2010; Kinney et al., 2012). Though structurally resembling icosahedral virus capsids, β-carboxysomes present a soft mechanical fingerprint (Faulkner et al., 2017).

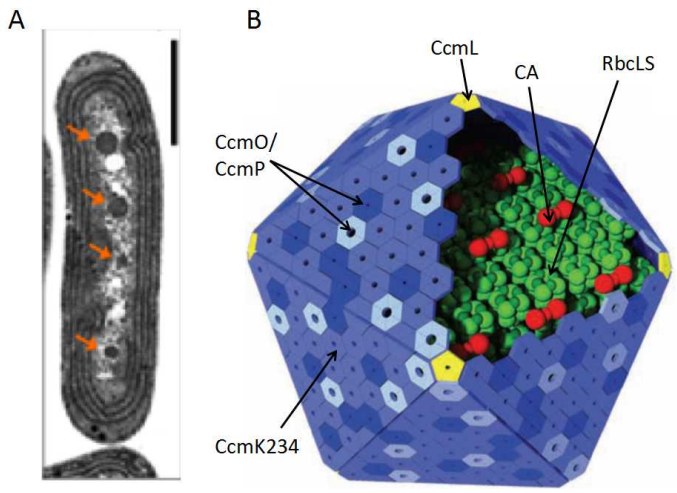

Figure 3-1. The physical organization of the β-carboxysomes.
A. Thin-section TEM image of *Synechococcus elongatus* PCC 7942 cells grow under ML. Arrows show carboxysomes with polyhedral shape (Scale bar: 1 μM) (Sun et al., 2016).
B. A model of the β-carboxysome structure as an icosahedron (Zarzycki et al., 2013).

The self-assembly, modularity and metabolic compartmentalisation and enhancement of BMCs make these bacterial organelles an ideal engineering objective (Frank et al., 2013). BMC shell proteins have been shown to self-assemble to form flat sheets or tubular structures (Pang et al., 2014; Noël et al., 2016; Sutter et al., 2016), and empty shells (Lassila et al., 2014; Cai et al., 2015; Held et al., 2016; Sutter et al., 2017). Expression of *ccmK1*, *ccmK2*, *ccmL*, and *ccmO* from the cyanobacterium *Halothece* sp. PCC 7418 in *Escherichia coli* could generate the synthetic β-carboxysome shells, ~ 25 nm in diameter (Cai et al., 2016). Likewise, empty BMC shells with the diameter of ~ 40 nm were obtained by expressing a synthetic operon of *Haliangium ochraceum* BMC shell genes (Lassila et al., 2014; Sutter et al., 2017). The empty BMC shells, without encapsulated enzymes, are notably smaller than the native BMCs. Furthermore, previous studies have demonstrated the possibilities of engineering entire PDU BMCs from *Citrobacter freundii* and EUT BMCs from *Salmonella enterica*

in *E. coli* (Parsons et al., 2008; Parsons et al., 2010; Choudhary et al., 2012). Expressing the α-carboxysome operon from a chemoautotroph *Halothiobacillus neapolitanus* has also led to the production of recombinant carboxysome-like structures with CO_2 fixation activity in *E. coli* (Bonacci et al., 2012) and in a Gram-positive bacterium *Corynebacterium glutamicum* (Baumgart et al., 2017). Apart from engineering BMCs in bacterial expression systems, efforts have been made to express β-carboxysome components in the chloroplasts of the model plant tobacco *Nicotiana benthamiana*. It was illustrated that transient expression of β-carboxysome proteins CcmK2, CcmM, CcmL, CcmO and CcmN could lead to the formation of carboxysome-like circular structures (Lin et al., 2014a). Moreover, Rubisco enzymes with a high carboxylation rate could be produced in tobacco chloroplasts by importing Syn7942 Rubisco-CcmM35 complexes to replace endogenous plant Rubisco (Lin et al., 2014b; Occhialini et al., 2016). Despite the recent advancement of carboxysome engineering, until now, heterologous production of entire functional β-carboxysome structures in any non-native hosts has not been reported, to our knowledge.

In this study, we generated a synthetic operon consisting of 12 β-carboxysome genes from Syn7942 and expressed the construct in *E. coli* to produce synthetic β-carboxysome-like structures with CO_2 fixation capacity. We further purified the synthetic β-carboxysomes and elucidated their structure, activity and interchangeability. This study improves our knowledge for engineering functional BMC structures in heterologous organisms and emphasises the necessity of optimizing the expression of carboxysome modules. It represents a step towards modulating BMC

structures and functions using synthetic biology, to produce new nanobioreactors and biomaterials with expanded biotechnological applications.

3.2 Results

3.2.1 Expression of Syn7942 β-carboxysomes in *Escherichia coli*

In contrast to the genes of α-carboxysomes that are mainly clustered together into a single operon, β-carboxysome genes typically appear in multiple dispersed gene clusters throughout the genome (Rae et al., 2013). In Syn7942, the twelve genes encoding β-carboxysomal proteins are distributed in five different chromosomal loci and the major carboxysome genes are structured in one locus containing the *ccmKLMNO* and *rbcLS* operons (Figure 3-2A). We generated a synthetic operon, by assembling all the twelve β-carboxysomal genes, to express β-carboxysomes heterologously. Additionally, the *ccmKLMNO* and *rbcLS* operons were assembled into a single synthetic operon, for producing β-carboxysome structures with a reduced number of genes to be expressed. These two operons, containing the native promoters of individual carboxysome genes, were then inserted into an *E. coli* expression vector pETM11 under a T7 promoter, to generate the pLFbC901 and pLFbC601 constructs in the host *E. coli* BL21DE3 (Figure 3-2A).

Expression of β-carboxysome components in *E. coli* was induced by IPTG (Figure 3-2D). The formation of β-carboxysome structures was verified by thin-section electron microscopy (Figure 3-2B). The carboxysome-like structures (~ 200 nm in

diameter) (Figure 3-2B) with a high internal protein density were observed in both recombinant *E. coli* cells harbouring pLFbC901 and pLFbC601 but are not present in the WT *E. coli* cells without carboxysome-expressing vectors (Figure 3-2B). The size of the carboxysome-like structures in pLFbC901 and pLFbC601 are similar (Figure 3-2B). Rubisco assays were performed to determine the *in vivo* carbon fixation activities of recombinant *E. coli* cells. Compared with WT BL21(DE3) cells which also treated with IPTG, both pLFbC601 and pLFbC901 exhibit high carbon fixation activities, suggesting the functioning of recombinant carboxysome-like structures in the heterologous expression (Figure 3-2C). In addition, the recombinant *E. coli* cells containing pLFbC901 present a higher carbon fixation activity than the pLFbC601 cells, likely due to the roles of CcaA, CcmK3K4, CcmP and RbcX in carboxysome structure and function. Because of the higher carbon fixation activity, we mainly focused on the pLFbC901 expression in the following study.

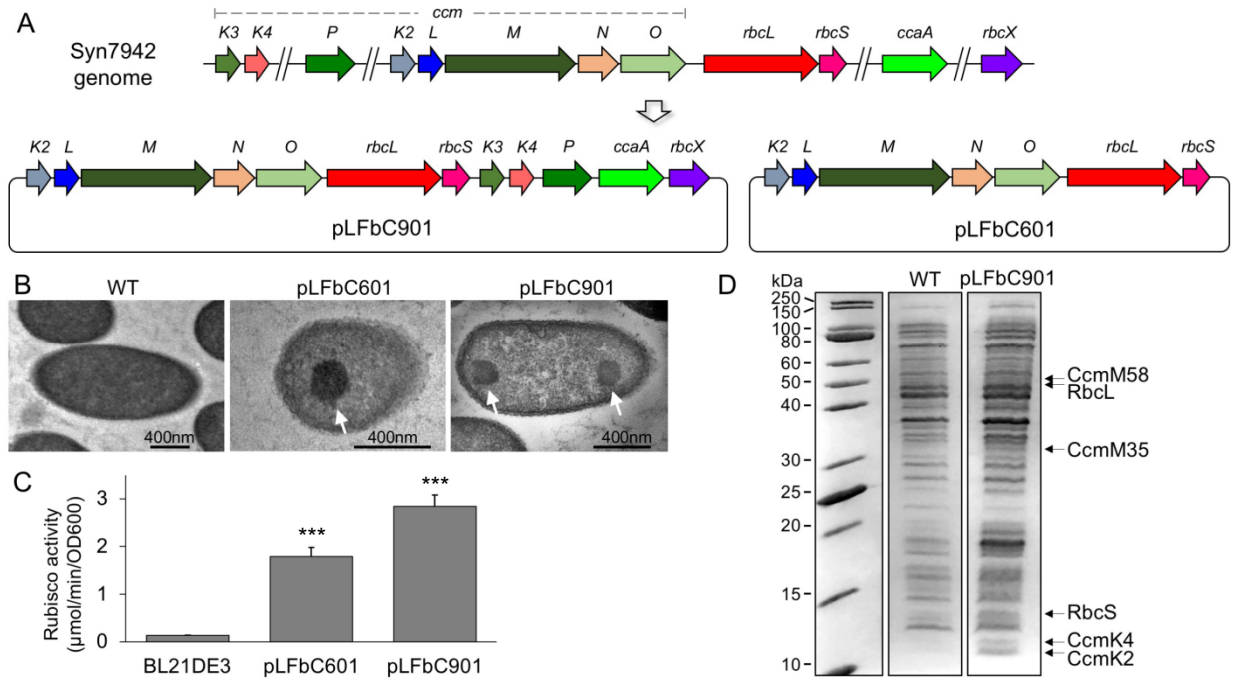

Figure 3-2. Synthetic operon design and heterologous expression of synthetic β-carboxysomes in *E. coli*.

A. Schematic representation of the molecular organization of the natural β-carboxysome operons spread in five loci within the Syn7942 genome and the synthetic β-carboxysome operons (pLFbC901 and pLFbC601) inserted in the *E. coli* expression vector pETM11. Locus tags are indicated.

B. Thin-section TEM images of *E. coli* WT and cells expressing β-carboxysome proteins using pLFbC601 and pLFbC901 plasmids. Arrows indicate β-carboxysome-like structures with polyhedral shapes observed in pLFbC601 and pLFbC901 *E. coli* cells.

C. *In vivo* carbon fixation activities of *E. coli* WT, pLFbC901 and pLFbC601 cells, indicating the CO_2-fixing functions of synthetic β-carboxysomes. A relatively higher CO_2-fixing activity of pLFbC901 *E. coli* cells compared to that of pLFbC601 *E. coli* cells was determined. Data are presented as mean ± SD, results come from 3 biological repeat. Statistical analysis implies the significant difference (***p<0.001, two-tailed t-test).

D. SDS-PAGE of the total cell extracts of the recombinant *E. coli* pLFbC901 and the control *E. coli* cells. Putative β-carboxysome proteins (RbcL, RbcS, CcmM, CcmK2, CcmK4) were identified based on their molecular weights.

CyanoBase	UniProt	Protein Description	Masco t score	Coverage (%)
Synpcc7942_1426	Q31NB3	Rubisco large chain RbcL	206	14
Synpcc7942_1427	Q31NB2	Rubisco small chain RbcS	178	42
Synpcc7942_1421	Q03511	CcmK2	155	40
Synpcc7942_1423	Q03513	CcmM	129	6
Synpcc7942_0285	Q31RK2	CcmK4	112	24

Table 3-1. Proteomic analysis of identified β-carboxysome proteins expressed by pLFbC901 in the whole *E. coli* cell extract.

To evaluate the assembly of β-carboxysomal proteins in E. coli, we generated another two plasmids. One (RbcL-eGFP-RbcS) contains the *rbcLS* operon with the enhanced green fluorescent protein (eGFP) gene fused at the 3' end of rbcL, which encodes the interior protein RbcL (Rubisco large chain) and the other (CcmK4-eGFP) comprises the gene of the shell protein CcmK4 fused with *egfp* at the 3' end. The two plasmids were transformed into *E. coli* strains, respectively, and were expressed in the absence or presence of pLFbC901 (Figure 3-3). *E. coli* WT cells were treated in the same conditions as the control (Figure 3-3, left). Without the expression of pLFbC901, the expressed RbcL-eGFP-RbcS and CcmK4-eGFP are evenly distributed in the cytoplasm, and no detectable protein aggregation was visualized under this induction

condition (Figure 3-3, middle). Co-expression of pLFbC901 with RbcL-eGFP-RbcS or CcmK4-eGFP can trigger the formation of protein assemblies (Figure 3-3, right), which shows the fluorescence spots exhibiting in the cytoplasm, and they also have a notable propensity to localized at the cell poles (Figure 3-3, right, arrows). These cell pole assemblies could be the inclusion bodies caused by the overexpression of RbcL or CcmK4, as has been discerned in the heterologous expression of PDU BMCs and α-carboxysomes (Parson et al., 2010; Bonacci et al., 2012). However, the cytoplasmic foci observed when co-expressing RbcL-eGFP-RbcS/CcmK4-eGFP and the recombinant β-carboxysome operon are reminiscent of the β-carboxysome structures observed in Syn7942 and may represent the β-carboxysome self-assemblies (Savage et al., 2010; Cameron et al., 2013).

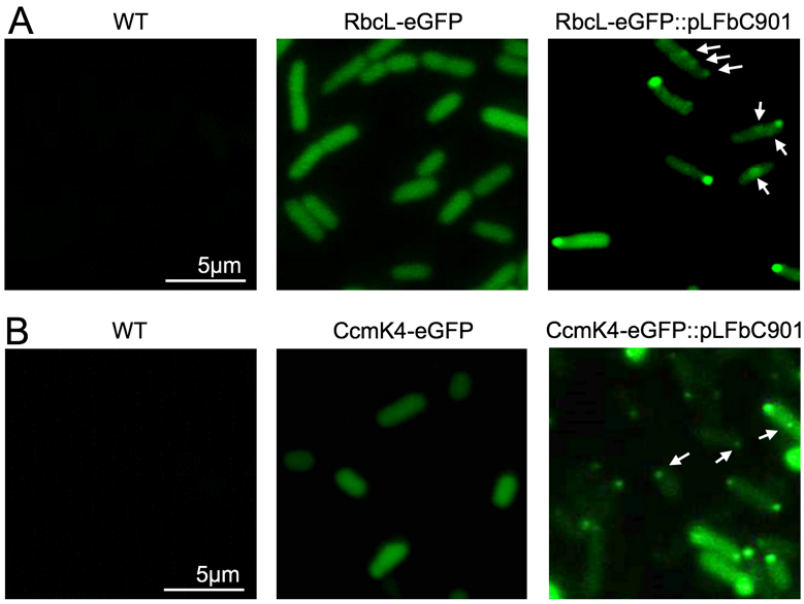

Figure 3-3. Fluorescence microscopy images of *E. coli* WT and cells expressing β-carboxysome components.

A. *E. coli* cells expressing RbcL-eGFP-RbcS in the absence or presence of the pLFbC901 vector. The blank *E. coli* cells were treated with the same conditions as the control.

B. *E. coli* cells expressing CcmK4-eGFP in the absence or presence of the pLFbC901 vector. The appearance of eGFP fluorescent puncta (arrows) indicates the assembly of β-carboxysome proteins in *E. coli*.

3.2.2 Purification and characterization of synthetic β-carboxysomes

Purification of recombinant β-carboxysomes was performed by sucrose density gradient centrifugation, following the induction of *E. coli* constructs (Figure 3-4A). We used the RbcL-eGFP-RbcS and pLFbC901 co-expression strain to facilitate the purification and detection of synthetic β-carboxysomes. The eGFP tagging has been used in isolating natural β-carboxysomes from Syn7942 (Faulkner et al., 2017). The majority of synthetic β-carboxysomes-like structures were detected in the 50% sucrose fractions by fluorescence imaging (Figure 3-4B). Rubisco assays of each β-carboxysome fraction also reveal that the 50% fraction presents the highest

Rubisco activity (Figure 3-4C). By contrast, native β-carboxysomes from Syn7942 were found in the 40% fraction (Faulkner et al., 2017). This is ascribed to the relatively larger size of synthetic β-carboxysomes-like structures (~ 200 nm, Figure 3-2B) in contrast to native β-carboxysomes (~ 150 nm)(Faulkner et al., 2017), which might because of several factors, like missing chaperones to help Rubisco assembly, missing Rubisco accumulating factors or different cytoplasmic conditions.

The same purification strategy is also applicable to synthetic β-carboxysomes without eGFP fusion and functional synthetic β-carboxysomes could be obtained in the 50% fraction, as confirmed by ^{14}C radiometric Rubisco assays (Figure 3-4D). It is worth noting that the yield of synthetic β-carboxysomes was low and the carboxysome proteins were not clearly identifiable in SDS-PAGE. However, immunoblot analysis using anti-RbcL and anti-CcmK2 revealed the presence of interior and shell proteins in isolated synthetic β-carboxysomes (Figure 3-4E).

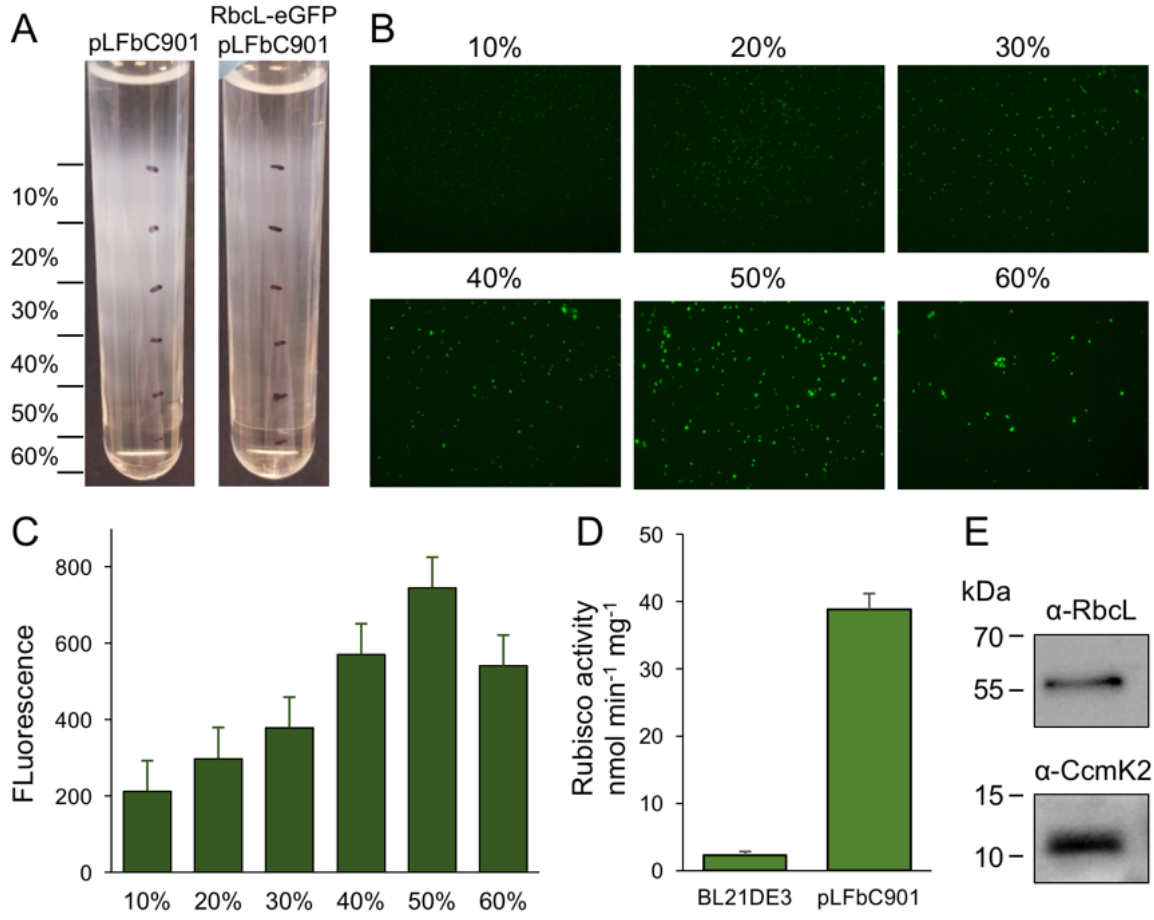

Figure 3-4. Purification and characterization of synthetic β-carboxysomes produced in *E. coli*.

A. Step sucrose gradient separation of synthetic β-carboxysomes with and without RbcL-eGFP-RbcS from pLFbC901.

B. Fluorescence detection of β-carboxysomes fused with eGFP in different sucrose fractions.

C. Fluorescence intensities of eGFP signals in different sucrose fractions, indicating the aggregation and abundance of β-carboxysome structures in each sucrose fraction.

D. Rubisco activities of β-carboxysomes in 50% sucrose fraction determined by ^{14}C radiometric assay.

E. Immunoblot analysis of 50% sucrose fraction using α-RbcL and α-CcmK2 antibodies, corroborating the presence of interior (RbcL) and shell proteins (CcmK).

Electron microscopy images of purified β-carboxysomes describe that synthetic β-carboxysomes-like structures are 200-300 nm in diameter (Figures 3-5A and 5B), slightly larger than native β-carboxysomes from Syn7942 (Faulkner et al., 2017). Moreover, although they do not exhibit manifestly a regular icosahedral shape and

symmetry, the particles still show the clear edges and vertices which is typical in carboxysome. Despite the possibility of being artifacts in sample preparation, the structural variations of native and synthetic β-carboxysomes-like structures may be a consequence of the change in protein abundance, ratio and organization, which requires further optimization.

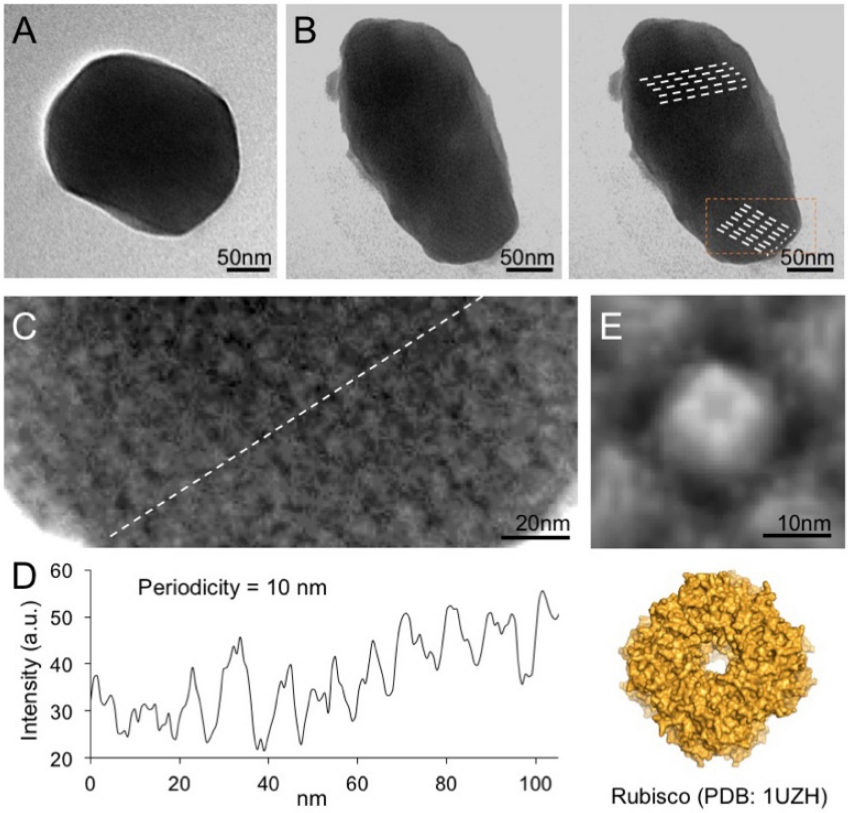

Figure 3-5. Negative-staining TEM images of isolated recombinant β-carboxysomes.

A. Representative β-carboxysome structures showing the polyhedral shape. The diameter of synthetic β-carboxysomes falls in the range of 200-300 nm.

B. Irregular β-carboxysome structures observed, implying the heterogeneity of recombinant β-carboxysomes. However, the densely-packed paracrystalline arrays of interior proteins could be discerned (dash lines).

C. Zoom-in inspection of the paracrystalline arrays depicted in (B, rectangle).

D. Cross-section profile analysis of the paracrystalline arrays illustrates the periodic arrangement of interior proteins and the center-to-center distance between neighboring proteins is ~10 nm.

E. Four-fold symmetrized correlation average TEM image of the interior protein (top) matches the atomic structure of Rubisco holoenzymes (bottom, PDB: 1UZH), indicating that the observed densely-packed interior particles are Rubisco enzymes.

Closer inspection of the electron micrographs reveals the highly-ordered paracrystalline arrays of interior proteins within the β-carboxysomal lumen (Figures 3-5B and 5C). Cross-section analysis by Plot Profile (ImageJ) illustrates that the

center-to-center distance of the interior proteins is about 10 nm (Figure 3-5D), the same as previous research (Faulker et al., 2017) . The specific protein arrangement observed in synthetic β-carboxysome-like particles is in good agreement with the Rubisco packing of the native β-carboxysome discriminated in previous studies (Kaneko et al., 2006; Faulkner et al., 2017). The individual proteins in the paracrystalline arrays depicted in the EM image were further analyzed using four-fold symmetrized cross-correlation single-particle averaging (Fechner et al., 2009; Liu et al., 2011d; Casella et al., 2017). The average protein structure matches well the atomic structure of Rubisco holoenzyme (Figure 3-5E). All of these results suggest that these proteins are Rubisco enzymes. Furthermore, because we can also detect the shell proteins from the isolated structure (Figure 3-4E), we suppose the isolated particles are the densely packed Rubisco and carboxysomes shell proteins assembled into carboxysome-like structures.

Furthermore, we studied the architecture of synthetic β-carboxysomes in solution using atomic force microscopy (AFM). The identification and structural integrity of eGFP-fused β-carboxysomes (RbcL-eGFP-RbcS) were verified by simultaneous AFM-fluorescence imaging (Figure 3-6). The single synthetic β-carboxysome-like structure is ~ 250 nm large and ~ 180 nm high, slightly larger than the native β-carboxysome and consistent with TEM results (Figures 3-2B and 3-2A) (Faulkner et a., 2017).

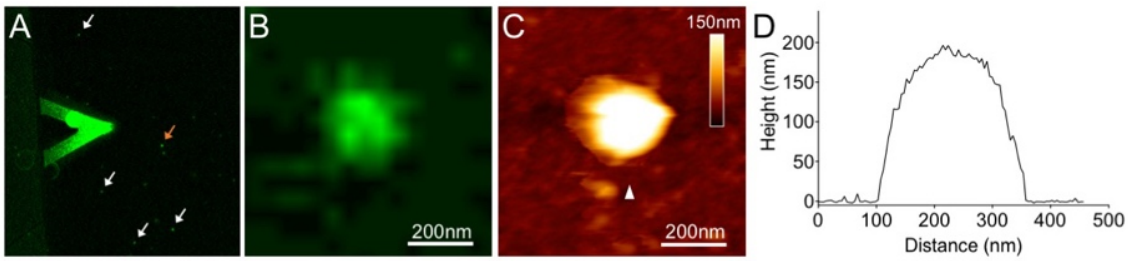

Figure 3-6. Combined confocal and AFM imaging of isolated synthetic β-carboxysomes fused with eGFP.

A. Merged microscopic image captured using a hybrid JPK AFM-Zeiss 880 confocal microscope. The triangular AFM cantilever is visible and arrows indicate individual β-carboxysome particles immobilized on mica surface.

B. Fluorescence image of a single recombinant β-carboxysome indicated by the orange arrow in A.

C. AFM topograph of the β-carboxysome captured simultaneously with the fluorescence image B. The combination of AFM-confocal fluorescence imaging ensures the identification of intact recombinant β-carboxysomes.

D. Height profile of the β-carboxysome structure, taken along the white arrowhead indicated in B.

3.2.3 Expressing synthetic β-carboxysomes in Syn7942

Our study shows that the designed synthetic carboxysome operon can produce functional β-carboxysomes in *E. coli*. To verify if the synthetic β-carboxysome operon can be expressed in Syn7942 cells to enhance carbon fixation, we inserted the synthetic β-carboxysome operon into a Syn7942 expression vector pAM2991 under the control of IPTG-inducible Ptrc promoter (Ivleva et al., 2005). The generated vector was then transformed into the RbcL-eGFP Syn7942 cells grown in low light (10 μE・m^{-2}・s^{-1}) and expressed in the presence of 100 μM IPTG (Figure 3-7A). The RbcL-eGFP Syn7942 cells express fluorescently tagged β-carboxysomes under endogenous genetic regulation (Sun et al., 2016). Confocal fluorescence images show that without IPTG induction, the Syn7942 cells harbour 1.7 ± 0.78

carboxysomes per cell (mean ± SD, *n* = 250, Figure 3-7B), consistent with the previous observation(Sun et al., 2016). Meanwhile, 100 μM IPTG induction didn't obviously affect the size and the phenotype of the Syn7942 cells and results in the expression of foreign β-carboxysome proteins which lead to a 50% increase in carboxysome number per cell (2.6 ± 1.2, mean ± SD, *n* = 350). The carboxysomes are spatially distributed along the long axis of the cell. Moreover, Rubisco assays elucidate a 13.7% increase in carbon fixation activity per cell, due to the elevated abundance of functional β-carboxysomes *in vivo* (Figure 3-7C). These results describe that the expression of synthetic β-carboxysome operon can trigger the formation of extra β-carboxysomes that are physiological positioned and functional in native hosts.

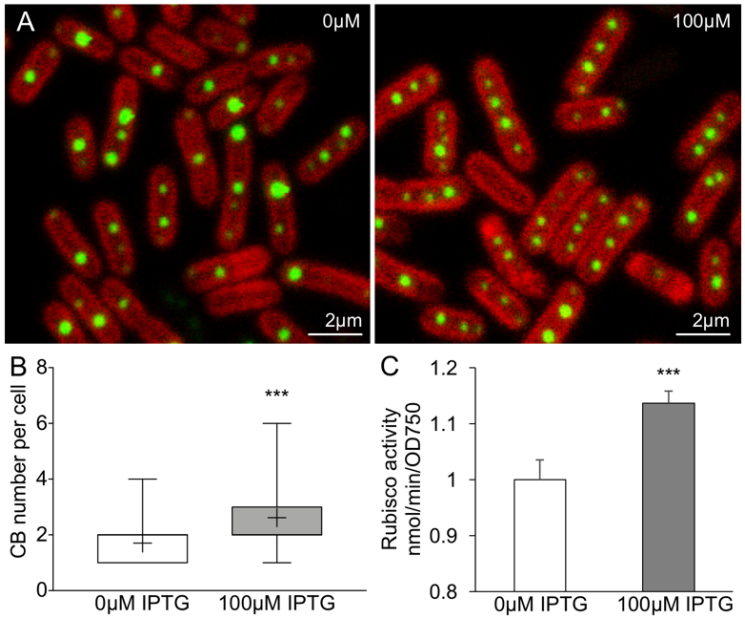

Figure 3-7. Enhancement of β-carboxysome formation in Syn7942 by the expression of pLFbC901.

A. Confocal microscopy images illustrating the β-carboxysome formation and abundance in Syn7942 cells induced by IPTG (left, no IPTG treatment; right, 100 μM IPTG). Chlorophyll fluorescence is shown in red and fluorescently tagged β-carboxysomes are discerned as green puncta.

B. The number of β-carboxysomes per cell increases from 1.7 (no IPTG induction) to 2.6 (100 μM IPTG induction). The whiskers show data from minimum to maximum, "+" represents the average. Statistical analysis implies the significant difference (*** p < 0.001, two-tailed t-test).

C. *In vivo* ^{14}C radiometric assays of Syn7942 cells induced by IPTG, illustrating a 13.7% increase in Rubisco activity resulted in 100 μM IPTG induction. Data are presented as mean ± SD. *** p < 0.001 (n = 3, two-tailed t-test).

3.2.4 Modularity and interchangeability of the β-carboxysome structure

Distinct BMCs are made of protein subunits belonging to a family of homologous proteins. We further evaluated the modularity of the β-carboxysome structure and interchangeability of building blocks from distinct types of BMCs. The Form 1A Rubisco large subunit CbbL of α-carboxysomes from *Halothiobacillus neapolitanus* was fused with eYFP at the C-terminus and expressed in RbcL-CFP Syn7942 cells. Confocal images reveal the explicit co-localization of CbbL-eYFP and RbcL-CFP

(Figure 3-8A), corroborating that α-carboxysome CbbL can be encapsulated in the β-carboxysome lumen that comprises Form IB Rubisco enzymes, probably due to the high sequence similarity between CbbL and RbcL (78%, Figure 3-9). Likewise, the α-carboxysome shell protein CsoS1A was fused with CFP at the C-terminus and expressed in CcmK4-eYFP Syn7942 cells. The notable co-localization of CsoS1A-CFP and CcmK4-eYFP implies the incorporation of α-carboxysome shell proteins into β-carboxysome structures (Figure 3-8B), consistent with the previous finding (Cai et al., 2015).

In addition, the PDU microcompartment shell protein PduA from *Salmonella enterica* serovar Typhimurium (*S. Typhimurium*) fused with superfolder GFP (sfGFP) was inserted in pAM2991 and expressed in Syn7942. PduA-sfGFP forms fluorescent puncta and presents a typical pattern of β-carboxysome distribution in the cytoplasm of Syn7942 cells (Savage et al., 2010; Cameron et al., 2013), indicative of the integration of PDU BMCs shell proteins into the β-carboxysome structure (Figure 3-8C).

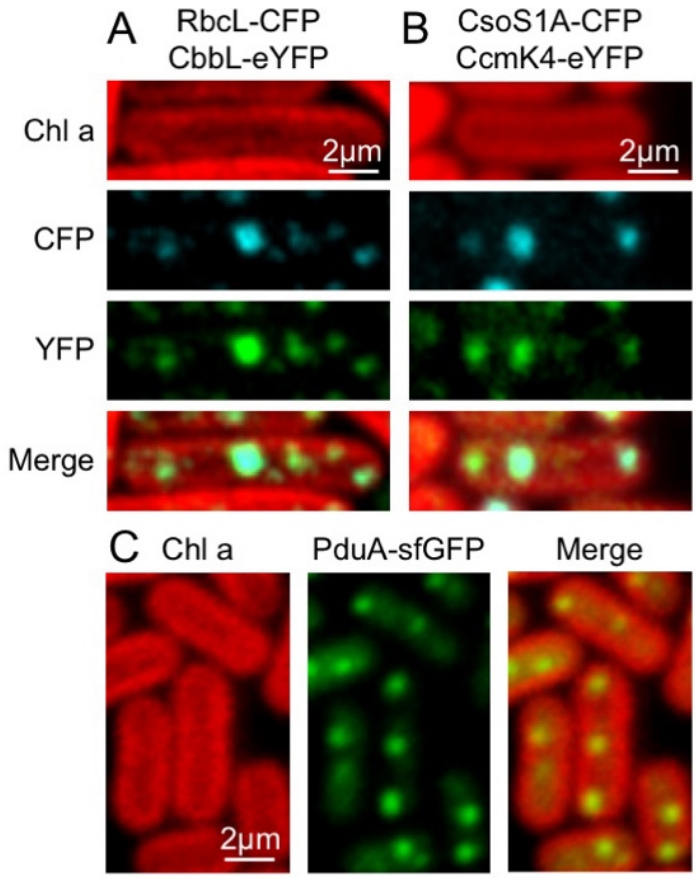

Figure 3-8. Interchangeability of BMC building blocks.

A. Co-localization of RbcL-CFP and CbbL-eYFP suggesting that α-carboxysome CbbL from *H. neapolitanus* can be encapsulated in the cyanobacterial β-carboxysome structures.

B. Co-localization of CsoS1A-CFP and CcmK4-eYFP implying the integration of α-carboxysome shell proteins from *H. neapolitanus* into the cyanobacterial β-carboxysome shell.

C. Evenly distributed fluorescence puncta of PduA-sfGFP along the longitude axis of Syn7942 cells indicating the incorporation of PDU shell proteins from *S. Typhimurium* in cyanobacterial β-carboxysomes.

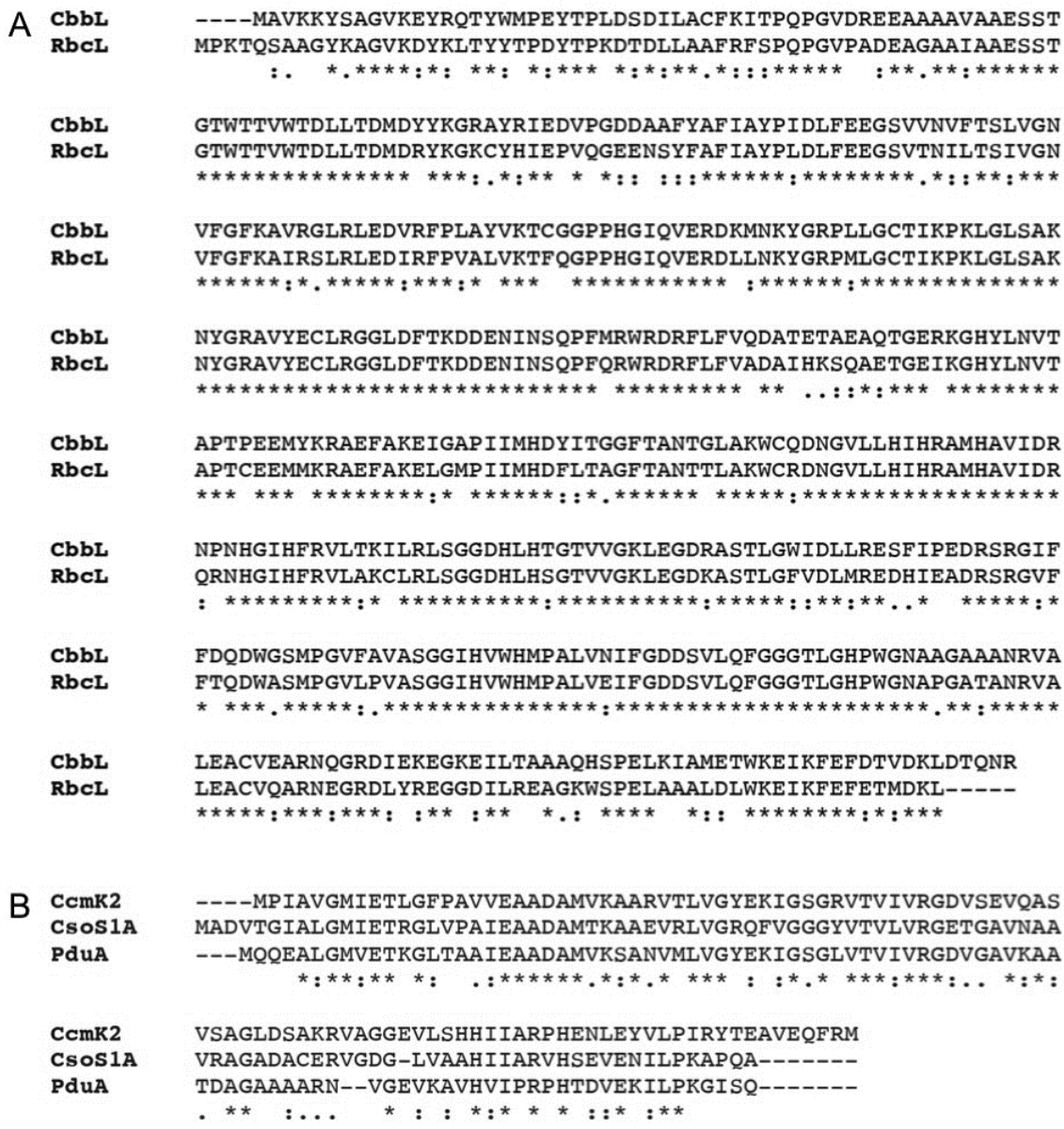

Figure 3-9. Sequence alignment of representative BMC proteins using Clustal Omega multiple sequence alignment program.

A. Sequence similarity (78%) between the Form 1B Rubisco large subunit CbbL from *H. neapolitanus* and the Form 1B Rubisco large subunit RbcL from Syn7942 (* fully conserved residue; : conservation between groups of strongly similar properties; . conservation between groups of weakly similar properties).

B. Sequence conservation of representative BMC shell hexamers, CcmK2 from Syn7942, CsoS1A from *H. neapolitanus* and PduA from *S. Typhimurium* (CsoS1A – CcmK2: 53%; PduA – CcmK2: 61%).

3.3 Discussion

The self-assembly and modularity features, as well as the specific role for CO_2 fixation, make carboxysomes a desirable engineering objective for metabolic enhancement and production of new nanomaterials. In this work, we show the construction of synthetic β-carboxysome operons and, to our knowledge, the first production of functional β-carboxysome-like structures in heterologous organisms. The structure and activity of synthetic β-carboxysomes, as well as the modularity of β-carboxysome structures, were characterized using a combination of molecular genetics, Rubisco assays, proteomic and immunoblot analysis, confocal, electron and atomic force microscopy.

The synthetic β-carboxysome operon was created to assemble 12 β-carboxysome genes. The recombinant β-carboxysomes, generated by the expression of the synthetic β-carboxysome operon, harbour densely packed Rubisco enzymes, CcmM and shell proteins (CcmK2, CcmK4), as confirmed by proteomic data (Table 3-1), immunoblot analysis (Figure 3-4E), and Rubisco assay (Figure 3-4D). Our results reveal that these proteins are the main building blocks necessary for β-carboxysome formation and CO_2 fixation. The “inside-out” model of β-carboxysome biogenesis suggests that the assembly of the β-carboxysome is initiated from the Rubisco-CcmM aggregation into procarboxysomes (Cameron et al., 2013), followed by the addition of CcmN which has a C-terminal encapsulation peptide (Kinney et al., 2012) to recruit shell proteins and result in the encapsulation of CcmK2 and CcmO. Seven out of the 12 β-carboxysome proteins were not detectable in proteomic analysis (Table 3-1).

This may suggest the low protein abundance in β-carboxysomes or the impeded protein expression (i.e. CcaA), probably due to the distant gene location from the T7 promoter. Nevertheless, the yield of synthetic β-carboxysomes after purification is low, confirming the overall expression of β-carboxysome proteins in the recombinant *E. coli* is restricted. Whether there are negative effects of the production of carboxysome-like structures on the growth of *E. coli* cells remains unknown. Moreover, the sizes of synthetic and native β-carboxysomes are relatively comparable (Figure 3-5), whereas the shape of synthetic β-carboxysomes is not well defined as the natural forms, likely as a result of improper content and ratio of expressed carboxysome proteins. All these highlights the necessity of optimising the expression constructs and conditions in future work as well as the β-carboxysome engineering in other hosts.

Despite the significance of Rubisco nucleation in shell encapsulation and as such β-carboxysome formation (Cameron et al., 2013), it has been demonstrated that synthetic β-carboxysome shell structures can be built by expressing the shell proteins CcmK1, CcmK2, CcmO, and CcmL (Cai et al., 2016). The synthetic β-carboxysome shells are about 25 nm in diameter, notably smaller than the native β-carboxysomes as well as synthetic β-carboxysomes produced in this study, suggesting the roles of interiors in determining the overall shape of β-carboxysomes.

The majority of cyanobacteria possess Form 1B Rubisco, the same as plants. However, compared with β-cyanobacteria Rubisco, plant Rubisco exhibits a relatively a low carboxylation activity but a high affinity for CO_2 against O_2. Improving the

catalytic activity of plant Rubisco has become a target of plant engineering for enhanced photosynthesis and capacity (Carmo-Silva et al., 2015). One possible strategy is to engineer functional cyanobacterial carboxysomes into plants and encapsulate plant Rubisco with the carboxysome shell that has the selective permeability to HCO_3^- and CO_2. Prior to this study, it has been demonstrated that expressing β-carboxysome shell proteins and cyanobacterial Form 1B Rubisco in tobacco chloroplasts is achievable (Lin et al., 2014a; Lin et al., 2014b; Occhialini et al., 2016). Here, we describe the heterologous expression of functional β-carboxysomes in *E. coli*, which holds promise for plant chloroplast engineering to construct entire functional carboxysome structures for enhanced plant photosynthesis and productivity.

It is potentially challenging to perform the multi-gene transformation in higher plants. The production of functional β-carboxysomes requires not only the expression of 12 carboxysome proteins and possibly other associated proteins, but also is likely to involve the manipulation of protein abundance and ratio. In a recent study, a chimeric protein peptide CcmC, including three Rubisco small subunit-like domains, one CcaA and the CcmN C-terminus encapsulation peptide, was designed to replace the endogenous β-carboxysome core (CcmM, CcmN, CcaA)(Gonzalez-Esquer et al., 2015). The reengineered β-carboxysomes were shown to support photosynthesis *in vivo*. In this study, we showed that the capacity of CO_2 fixation can be obtained using a simplified carboxysome-expressing vector pLFbC601, which contains only seven β-carboxysome genes in the *ccmKLMNO* and *rbcLS* operons. This may represent a

promising engineering objective to produce functional carboxysomes in higher plants towards with a simplified composition. Optimization of the pLFbC601 expression to build functional CO_2-fixing modules awaits further investigation, i.e. enhancing and controlling the expression of individual genes, adjusting protein ratio and integrating chloroplast carbonic anhydrases.

Despite the functional diversity, all BMCs share a common structural feature: an icosahedral protein shell encapsulating internal catalytic enzymes (Bobik et al., 2015) The shell proteins have conserved sequences and form typical BMC domains: the PF00936 domain for hexameric shell proteins, the PF03319 domain for shell pentamers, and the PF00936 domain for trimeric shell proteins (Kerfeld et al., 2005; Tanaka et al., 2008; Klein et al., 2009). We demonstrate that the α-carboxysome shell protein CsoS1A from *Halothiobacillus neapolitanus*, the PDU shell hexameric protein PduA from *S. Typhimurium*, as well as the Form 1A Rubisco large subunit CbbL from *Halothiobacillus neapolitanus*, can be integrated into the native β-carboxysomes in Syn7942 cells. This is likely due to on their sequence conservation and structural similarities relative to CcmK2 and RbcL, respectively. Our results shed lights on the inherent modularity and interchangeability of a range of BMC building blocks, which lay the foundation for generating chimeric BMC structures in a predictive manner, to manipulate the metabolic activities of BMCs in native and heterologous hosts and engineer BMC-like nanoreactors for new functions. It will expand the potential applications of BMC-like structures in biotechnology.

3.4 References

Axen, S. D., O. Erbilgin, and C. A. Kerfeld. 2014. 'A taxonomy of bacterial microcompartment loci constructed by a novel scoring method', *PLoS Comput Biol*, 10: e1003898.

Baker, S. H., S. Jin, H. C. Aldrich, G. T. Howard, and J. M. Shively. 1998. 'Insertion mutation of the form I cbbL gene encoding ribulose bisphosphate carboxylase/oxygenase (RuBisCO) in Thiobacillus neapolitanus results in expression of form II RuBisCO, loss of carboxysomes, and an increased CO_2 requirement for growth', *J Bacteriol*, 180: 4133-9.

Baumgart, M., I. Huber, I. Abdollahzadeh, T. Gensch, and J. Frunzke. 2017. 'Heterologous expression of the *Halothiobacillus neapolitanus* carboxysomal gene cluster in *Corynebacterium glutamicum*', *J Biotechnol*, 258: 126-35.

Bobik, T. A., B. P. Lehman, and T. O. Yeates. 2015. 'Bacterial microcompartments: widespread prokaryotic organelles for isolation and optimization of metabolic pathways', *Mol Microbiol*, 98: 193-207.

Bonacci, W., P. K. Teng, B. Afonso, H. Niederholtmeyer, P. Grob, P. A. Silver, and D. F. Savage. 2012. 'Modularity of a carbon-fixing protein organelle', *Proc Natl Acad Sci U S A*, 109: 478-83.

Cai, F., S. L. Bernstein, S. C. Wilson, and C. A. Kerfeld. 2016. 'Production and characterization of synthetic carboxysome shells with incorporated luminal proteins', *Plant Physiol*, 170: 1868-77.

Cai, F., B. B. Menon, G. C. Cannon, K. J. Curry, J. M. Shively, and S. Heinhorst. 2009. 'The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier', *PLoS One*, 4: e7521.

Cai, F., M. Sutter, S. L. Bernstein, J. N. Kinney, and C. A. Kerfeld. 2015. 'Engineering bacterial microcompartment shells: chimeric shell proteins and chimeric carboxysome shells', *ACS Synth Biol*, 4: 444-53.

Cameron, J. C., S. C. Wilson, S. L. Bernstein, and C. A. Kerfeld. 2013. 'Biogenesis of a bacterial organelle: the carboxysome assembly pathway', *Cell*, 155: 1131-40.

Carmo-Silva, E., J. C. Scales, P. J. Madgwick, and M. A. Parry. 2015. 'Optimizing Rubisco and its regulation for greater resource use efficiency', *Plant Cell Environ*, 38: 1817-32.

Casella, S., F. Huang, D. Mason, G. Y. ZHAO, G. N. Johnson, C. W. Mullineaux, and L. N. Liu. 2017. 'Dissecting the native architecture and dynamics of cyanobacterial photosynthetic machinery', *Mol Plant*, 10: 1434-48.

Chen, A. H., B. Afonso, P. A. Silver, and D. F. Savage. 2012. 'Spatial and temporal organization of chromosome duplication and segregation in the cyanobacterium Synechococcus elongatus PCC 7942', *PLoS One*, 7: e47837.

Choudhary, S., M. B. Quin, M. A. Sanders, E. T. Johnson, and C. Schmidt-Dannert. 2012. 'Engineered protein nano-compartments for targeted enzyme localization', *PLoS One*, 7: e33342.

Chowdhury, C., S. Sinha, S. Chun, T. O. Yeates, and T. A. Bobik. 2014. 'Diverse bacterial microcompartment organelles', *Microbiol Mol Biol Rev*, 78: 438-68.

Faulkner, Matthew, Jorge Rodriguez-Ramos, Gregory F. Dykes, Sian V. Owen, Selene Casella, Deborah M. Simpson, Robert J. Beynon, and Lu-Ning Liu. 2017. 'Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes', *Nanoscale*, 9: 10662–73.

Fechner, P., T. Boudier, S. Mangenot, S. Jaroslawski, J. N. Sturgis, and S. Scheuring. 2009. 'Structural information, resolution, and noise in high-resolution atomic force microscopy topographs', *Biophys J*, 96: 3822-31.

Frank, S., A. D. Lawrence, M. B. Prentice, and M. J. Warren. 2013. 'Bacterial microcompartments moving into a synthetic biological world', *J Biotechnol*, 163: 273-9.

Giessen, T. W., and P. A. Silver. 2016. 'Encapsulation as a strategy for the design of biological compartmentalization', *J Mol Biol*, 428: 916-27.

Gonzalez-Esquer, C. R., T. B. Shubitowski, and C. A. Kerfeld. 2015. 'Streamlined construction of the cyanobacterial CO_2-fixing organelle via protein domain fusions for use in plant synthetic biology', *Plant Cell*, 27: 2637-44.

Hammer, S. K., and J. L. Avalos. 2017. 'Harnessing yeast organelles for metabolic engineering', *Nat Chem Biol*, 13: 823-32.

Held, M., A. Kolb, S. Perdue, S. Y. Hsu, S. E. Bloch, M. B. Quin, and C. Schmidt-Dannert. 2016. 'Engineering formation of multiple recombinant Eut protein nanocompartments in *E. coli*', *Sci Rep*, 6: 24359.

Iancu, C. V., D. M. Morris, Z. Dou, S. Heinhorst, G. C. Cannon, and G. J. Jensen. 2010. 'Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells', *J Mol Biol*, 396: 105-17.

Ivleva, N. B., M. R. Bramlett, P. A. Lindahl, and S. S. Golden. 2005. 'LdpA: a component of the circadian clock senses redox state of the cell', *EMBO J*, 24: 1202-10.

Kaneko, Y., R. Danev, K. Nagayama, and H. Nakamoto. 2006. 'Intact carboxysomes in a cyanobacterial cell visualized by hilbert differential contrast transmission electron microscopy', *J Bacteriol*, 188: 805-8.

Kerfeld, C. A., and O. Erbilgin. 2015. 'Bacterial microcompartments and the modular construction of microbial metabolism', *Trends Microbiol*, 23: 22-34.

Kerfeld, Cheryl A., Michael R. Sawaya, Shiho Tanaka, Chau V. Nguyen, Martin Phillips, Morgan Beeby, and Todd O. Yeates. 2005. 'Protein structures forming the shell of primitive bacterial organelles', *Science*, 309: 936-38.

Kinney, J. N., A. Salmeen, F. Cai, and C. A. Kerfeld. 2012. 'Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly', *J Biol Chem*, 287: 17729-36.

Klein, M. G., P. Zwart, S. C. Bagby, F. Cai, S. W. Chisholm, S. Heinhorst, G. C. Cannon, and C. A. Kerfeld. 2009. 'Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport', *J Mol Biol*, 392: 319-33.

Lassila, J. K., S. L. Bernstein, J. N. Kinney, S. D. Axen, and C. A. Kerfeld. 2014. 'Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function', *J Mol Biol*, 426: 2217-28.

Lin, M. T., A. Occhialini, P. J. Andralojc, J. Devonshire, K. M. Hines, M. A. Parry, and M. R. Hanson. 2014. 'β-Carboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts', *Plant J*, 79: 1-12.

Lin, Myat T., Alessandro Occhialini, P. John Andralojc, Martin A. J. Parry, and Maureen R. Hanson. 2014. 'A faster Rubisco with potential to increase photosynthesis in crops', *Nature*, 513: 547-50.

Liu, L. N. 2016. 'Distribution and dynamics of electron transport complexes in cyanobacterial thylakoid membranes', *Biochim Biophys Acta*, 1857: 256-65.

Liu, L. N., J. N. Sturgis, and S. Scheuring. 2011. 'Native architecture of the photosynthetic membrane from *Rhodobacter veldkampii*', *J Struct Biol*, 173: 138-45.

Long, B. M., M. R. Badger, S. M. Whitney, and G. D. Price. 2007. 'Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA', *J Biol Chem*, 282: 29323-35.

Long, B. M., B. D. Rae, M. R. Badger, and G. D. Price. 2011. 'Over-expression of the beta-carboxysomal CcmM protein in Synechococcus PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content', *Photosynth Res*, 109: 33-45.

Long, B. M., L. Tucker, M. R. Badger, and G. D. Price. 2010. 'Functional cyanobacterial beta-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein', *Plant Physiol*, 153: 285-93.

Menon, B. B., Z. Dou, S. Heinhorst, J. M. Shively, and G. C. Cannon. 2008. 'Halothiobacillus neapolitanus carboxysomes sequester heterologous and chimeric RubisCO species', *PLoS One*, 3: e3570.

Noël, Christopher R., Fei Cai, and Cheryl A. Kerfeld. 2016. 'Purification and characterization of protein nanotubes assembled from a single bacterial microcompartment shell subunit', *Advanced Materials Interfaces*, 3: 1500295-n/a.

Occhialini, A., M. T. Lin, P. J. Andralojc, M. R. Hanson, and M. A. Parry. 2016. 'Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2', *Plant J*, 85: 148-60.

Pang, A., S. Frank, I. Brown, M. J. Warren, and R. W. Pickersgill. 2014. 'Structural insights into higher order assembly and function of the bacterial microcompartment protein PduA', *J Biol Chem*, 289: 22377-84.

Parsons, J. B., S. D. Dinesh, E. Deery, H. K. Leech, A. A. Brindley, D. Heldt, S. Frank, C. M. Smales, H. Lunsdorf, A. Rambach, M. H. Gass, A. Bleloch, K. J. McClean, A. W. Munro, S. E. Rigby, M. J. Warren, and M. B. Prentice. 2008. 'Biochemical and structural insights into bacterial organelle form and biogenesis', *J Biol Chem*, 283: 14366-75.

Parsons, J. B., S. Frank, D. Bhella, M. Liang, M. B. Prentice, D. P. Mulvihill, and M. J. Warren. 2010. 'Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement', *Mol Cell*, 38: 305-15.

Peña, K. L., S. E. Castel, C. de Araujo, G. S. Espie, and M. S. Kimber. 2010. 'Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM', *Proc Natl Acad Sci U S A*, 107: 2455-60.

Rae, B. D., B. M. Long, M. R. Badger, and G. D. Price. 2012. 'Structural determinants of the outer shell of beta-carboxysomes in *Synechococcus elongatus* PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL', *PLoS One*, 7: e43871.

Rae, B. D., B. M. Long, M. R. Badger and G. D. Price. 2013. "Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO2 fixation in cyanobacteria and some proteobacteria." *Microbiol Mol Biol Rev* 77(3): 357-379.

Savage, D. F., B. Afonso, A. H. Chen, and P. A. Silver. 2010. 'Spatially ordered dynamics of the bacterial carbon fixation machinery', *Science*, 327: 1258-61.

Shively, J. M., F. Ball, D. H. Brown, and R. E. Saunders. 1973. 'Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*', *Science*, 182: 584-6.

Sun, Y., S. Casella, Y. Fang, F. Huang, M. Faulkner, S. Barrett, and L.N. Liu. 2016. 'Light modulates the biosynthesis and organization of cyanobacterial carbon fixation machinery through photosynthetic electron flow', *Plant Physiol*, 171: 530-41.

Sutter, M., B. Greber, C. Aussignargues, and C. A. Kerfeld. 2017. 'Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell', *Science*, 356: 1293-97.

Sutter, Markus, Matthew Faulkner, Clément Aussignargues, Bradley C Paasch, Steve Barrett, Cheryl A Kerfeld, and Lu-Ning Liu. 2016. 'Visualization of bacterial microcompartment facet assembly using high-speed atomic force microscopy', *Nano Letters*, 16: 1590-95.

Tanaka, Shiho, Cheryl A. Kerfeld, Michael R. Sawaya, Fei Cai, Sabine Heinhorst, Gordon C. Cannon, and Todd O. Yeates. 2008. 'Atomic-level models of the bacterial carboxysome shell', *Science*, 319: 1083-86.

Zarzycki, J., S. D. Axen, J. N. Kinney and C. A. Kerfeld. 2013. "Cyanobacterial-based approaches to improving photosynthesis in plants." *J Exp Bot* 64(3): 787-798.

Chapter 4

Synthetic engineering of α-carboxysomes in *E. coli*

4.1 Introduction

The oceanic cyanobacteria contain α-carboxysomes with Rubisco Form IA, known as α-cyanobacteria. The CCM of α-cyanobacteria is distinct from that of β-cyanobacteria regarding Ci uptake pathway, genetic architecture, and shell composition. The chemolithoautotrophic sulfur-oxidizing bacterium *Halothiobacillus neapolitanus* were pursued in order to investigate the roles and the assembly of α-carboxysome as the model system (Figure 4-1A) (Badger et al., 2003; Iancu et al., 2010). The α-carboxysomes of *H. neapolitanus* are composed of nine proteins from a single genomic locus, the *cso* operon, and a pseudohexameric shell protein CsoS1D. The α-carboxysome comprises Rubisco that consists of two subunits (large subunit CbbL and small subunit CbbS), the major shell protein CsoS1A/B/C, minor shell proteins CsoS4A/B and CsoS1D, shell-associated protein CsoS2, and carboxysomal CA enzyme CsoS3 (Figure 4-1B, 1C) (Rae et al., 2013; Klein et al., 2009; Roberts et al., 2012).

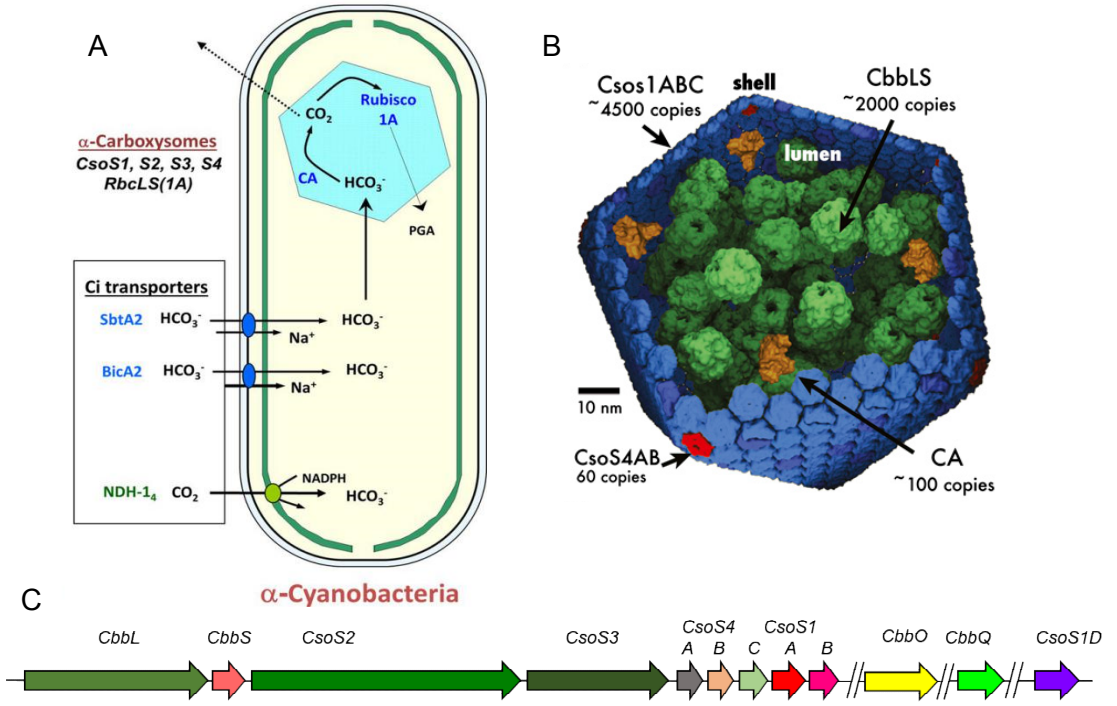

Figure 4-1. Physical and genetic organization of the α-cyanobacteria.

A. The general CCM characteristics and Ci transporters in α-cyanobacteria (Rae et al., 2013).

B. A partial model of the α-carboxysome structure (Bonacci et al., 2012).

C. Genomic organization of the cso operon and Rubisco activase in *H. neapolitanus*.

The Form 1A Rubisco inside α-carboxysomes is also composed of eight large and eight small subunits as Form 1B Rubisco (Badger et al., 2008). However, the electrostatic potential maps show that the Form 1A RubisCO has a much more negatively charged surface than that of the Form 1B enzymes (Zarzycki et al., 2013). It is thought that the small subunits of Rubisco are responsible for Rubisco encapsulation into carboxysomes by interacting with CsoS2 (Holthuijzen et al., 1986; Kelly et al., 2000).

By using Cryo-electron microscopy, the purified α-carboxysomes from *H. neapolitanus* (Schmid et al., 2006) and *Synechococcus* WH8102 (Iancu et al., 2007) were found in a range of sizes which might depend on different arrangements of shell

proteins. Inside carboxysomes, Rubisco aggregations were preferably arranged into the concentric layers and located in the layer adjacent to the shell (Iancu et al., 2010; Schmid et al., 2007; Iancu et al., 2007).

CsoS1A, CsoS1B, and CsoS1C are small major shell proteins that form the shell of the α-carboxysome. These three proteins are highly identical to each other in amino acid sequence, except for a short C-terminal extension in CsoS1B (Tsai et al., 2007; Cannon et al., 2001; Cannon et al., 2003; English et al., 1994), but their precise ratio is still unknown. The crystal structure of CsoS1A was reported as hexameric units which can pack tightly to form a molecular layer. Sulfate ion, which having a similar size and can be considered an analog for bicarbonate, was found in the pore of the hexamer, which indicated the negatively charged metabolites such as bicarbonate could cross the shell through the pore (Tsai et al., 2007).

Two paralogs CsoS4A and CsoS4B, homologous to CcmL from the β-carboxysome appear as pentamers located at the vertices of the carboxysome shell (Tanaka et al, 2008). However, although the absence of CsoS4A and CsoS4B together could result in some elongated carboxysomes, most of the carboxysomes still formed normally *in vivo* and *in vitro*, contrary to the prediction of the current α-carboxysome shell model (Cai et al., 2009; Price et al., 1993; Wang et al., 2014; Price et al., 1989). Further experiment showed that the CsoS4 protein is vital for CO_2 leakage and provide the catalytic advantage to RuBisCO derives (Cai et al., 2009).

CsoS2 is the carboxysome shell-associated protein and vital for α-carboxysome

biogenesis (Gonzales et al., 2005). It was thought to be the third most abundant carboxysome protein in purified carboxysomes and can interact with shell proteins and Rubisco (Gonzales et al., 2005; Baker et al., 1999; Cai et al., 2015). In *H. neapolitanus*, the yeast two-hybrid screen showed that CsoS2 protein strongly interacts with CsoS4B, CsoS1C, and CbbS (Cai et al., 2015). The primary structure of CsoS2 showed three regions, N-terminus enriched in α-helices (aa 1~259), the middle composed of β-strands (aa 260~603), and the conserved C-terminus (aa 604~869) (Cai et al., 2015). The *csoS2* gene can produce a full-length protein CsoS2B (130 kDa) and a shorter protein CsoS2A (92 kDa) in *H. neapolitanus*. These two proteins have a 1:1 stoichiometric expression level and share the same N-terminus while the C-terminus of CsoS2A was truncated, because of a ribosomal frameshifting cis elements (Chaijarasphong et al., 2016). Furthermore, intact carboxysomes can be formed with the expression of CsoS2B along, not CsoS2A, indicating that the N-terminus and the middle region of CsoS2 (CsoS2A/B) organize Rubisco in the core of the lumen while C-terminus of CsoS2B anchors them to the shell (Figure 4-2) (Chaijarasphong et al., 2016).

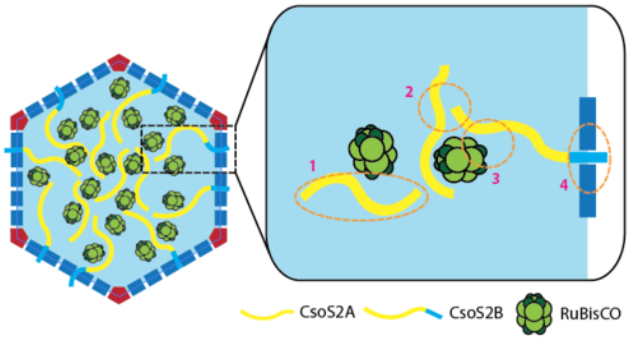

Figure 4-2. Model of α-carboxysome assembly.
Differential location and function of different CsoS2 isoforms in α-carboxysome assembly (Chaijarasphong et al., 2016).

CsoS3, also called CsoSCA, is a novel carbonic anhydrase, constitutes a subclass of β-carbonic anhydrases (Sawaya et al., 2006). CsoS3 forms a 57kDa homodimer and contains only one active site whereas other β-CAs contain two (Sawaya et al., 2006). This active site was thought to bind one Zn^{2+} per monomer which is important for CA activity (Heinhorst et al., 2006). Knockout of CsoSCA does not affect carboxysome assembly (Dou et al., 2008), suggesting that it is a shell-associated component, but it is associated with the carboxysome shell so tightly that it can only be released with denaturants (Heinhorst et al., 2006). CsoS3 is also inactivated under reducing condition (reducing agent to break the disulfidebonds, 10mM TCEP or 10mM DTT) which is similar to the CA from β-carboxysomes (So et al., 2004; Pena et al, 2010). The lower apparent catalytic rate of CsoSCA in intact α-carboxysomes than in ruptured carboxysomes indicated that the carboxysome protein shell acts as a barrier to capture CO_2 produced by CsoS3 (Heinhorst et al., 2006).

CsoS1D is another component of the carboxysome shell. The crystal structure of the CsoS1D protein from *Prochlorococcus* strain MED4 exhibits pseudohexamer block for

incorporation into the carboxysome shell which is constitute of trimers (Klein et al., 2009). It contains an unusually large pore with distinct open and closed conformations. This structure might be a gate to control the transit of RuBP (Roberts et al., 2012). Furthermore, introducing CsoS1D in the α-carboxysome engineering could modify the structural stability and the Rubisco activity of the heterologous synthetic carboxysomes (Bonacci et al., 2012).

Partially assembled shell proteins with attached Rubisco was observed by cryo-electron tomography, suggesting that the shell assembling and Rubisco aggregation is concomitant in α-carboxysomes (Kerfeld et al., 2016; Iancu et al., 2010). Moreover, the α-carboxysome shell can still stably assemble in the Rubisco deletion mutants and exhibit normal architecture and shape, confirming that the shell biogenesis and Rubisco sequestration are two independent processes (Menon et al., 2008).

To catalytically active Rubisco, the first step must be non-substrate CO_2 carbamylate Rubisco at the active-site lysine, followed by Mg_2^+ binding. However, premature binding to the RuBP prior to CO_2 binding makes Rubisco become an inactive complex, blocking its ability to react with either CO_2 or O_2 (Tsai et al., 2015; Bhat et al., 2017; Lorimer et al., 1976; Jordan et al., 1983; Mueller-Cajar et al., 2011). Furthermore, during catalytic reactions, Rubisco can also be inhibited by misfire products, such as xylulose-1,5-bisphosphate (XuBP), 2,3-pentodiulose-1,5-bisphosphate (PDBP) (Rubisco regulation a role for inhibitors), and night-time inhibitor 2-carboxy-D-arabinitol-1-phosphate (CA1P) (Bhat et al., 2017; Andralojc et al., 2012;

Parry et al., 2008).

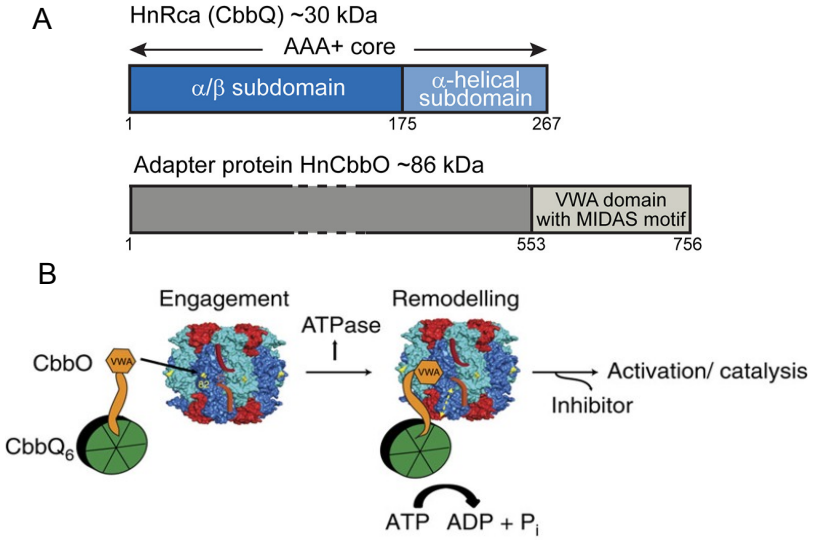

Figure 4-3. The prokaryotic Rubisco activase of the green-type Form IA Rubisco.

A. The domain structure of CbbQ from *H. neapolitanus* and its adapter protein CbbO (Bhat et al., 2017).

B. Model of the function of CbbQO for Form 1A Rubisco (Tsai et al., 2015).

Rubisco activase were thought to remodel rubisco complex to release inhibitor from inactive Rubisco (Tsai, et al., 2015; Bhat et al., 2017). They belong to the AAA^+ (ATPases associated with various cellular activities) protein superfamily as ATP-hydrolysing enzyme and contain a structurally conserved ATPase domain referred as AAA modules, consisting of an N-terminal P-loop with α/β fold, a nucleotide binding domain and a C-terminal α-helical domain (Teru et al., 2001). AAA^+ protein can form ring-shaped hexamers and are functionally active in the presence of ATP and RuBP. Rca was found in green algae, cyanobacteria and all plants which contain Rubisco Form 1B (Bhat et al., 2017), CbbX from red algae containing Rubisco Form IC and ID (Mueller-Cajar et al., 2011) and CbbQO complex from

chemoautotrophic bacteria containing Rubisco Form 1A (Figure 4-3) (Tsai, et al., 2015).

The sequences of *cbbQ* gene are highly conserved and they encode a ~30 kDa single AAA^+ subunit belonging to MoxR group. MoxR proteins often associated with proteins containing von Willebrand factor A (VWA) domain encoded by a downstream gene (Wong et al., 2012; Tsai, et al., 2015; Bhat et al., 2017). The crystal structure of CbbQ revealed an N-terminal P-loop NTPase domain, and two CbbQ molecules per unit form a ring-shaped hexamer which is typical for AAA^+-ATPase. The C-terminus of CbbQ is structurally unique (Figure 4-3A), this domain was mainly found on one surface of the hexamer, structurally filling the inter-monomer gaps and forming the concave face of the hexamer. This unique and high conserved surface might facilitate the interaction with substrates or other proteins (Sutter et al., 2015). Furthermore, CbbQ strongly interacts with CsoS1, indicating that it is associated with the carboxysome shell (Sutter et al., 2015).

CbbO has a C-terminal VWA domain with a metal-ion-dependent adhesion site (MIDAS) (Figure 4-3A). This site can be involved in protein interactions via a metal cation (Xiong et al., 2002; Santelli et al., 2004). Mutagenesis showed that MIDAS could interact with Rubisco large subunit (Tsai et al., 2015; Mueller-Cajar et al., 2011). These data support a model that involves CbbQ and CbbQ forming a bipartite complex comprising hexameric CbbQ activase, and CbbO as an adapter between the activase and Rubisco by its VWA domain (Figure 4-3B) (Tsai et al., 2015). Previous research showed that co-expression of CbbQ and/or CbbQ with Rubisco can increase

the rubisco activity when Rubisco is heterologously expressed in *E. coli* (Hayashi et al., 1997; Hayashi et al., 1999).

The functional α-carboxysomes have been synthesized in *E. coli* and *Corynebacterium glutamicum* with Rubisco activity (Bonacci et al., 2012; Baumgart, et al., 2017), and empty carboxysome shells lacking RuBisCO can be formed in *H. neapolitanus* (Menon et al., 2008). These results emphasized α-carboxysomes as orthogonal modules for synthetic biology approaches and highlighted the possibility of engineering and modulating α-carboxysomes for not only studying the characteristics of carboxysome biosynthesis but also engineering metabolic organelles in heterologous organisms. Introducing proteins which might help the function and assembly of carboxysomes, such as GroEL/ES, Rubisco assembly factor (RAF) and Rubisco activase, will help us understand their physiological roles. Given that functional α-carboxysomes can be synthetised in other hosts, proper Rubisco assembly appears not affected without GroEL/ES and RAF. It is not certain whether Rubisco activase is critical for the assembly and function of carboxysomes. In this study, I introduced the Rubisco activase CbbQ complex *H. neapolitanus* to heterologously express α-carboxysomes, which led to an enhanced carbon-fixation activity of α-synthetic carboxysomes.

4.2 Results

4.2.1 Expression of α-carboxysomes in *E. coli*

The pHnCB plasmid (Figure 4-4A) including the α-carboxysome *cso* operon was expressed in *E. coli* and results in the production of synthetic α-carboxysomes with well formed structure, expected molecular components, and able to fixing CO2 in vivo and in vitro (Bonacci et al., 2012). Furthermore, the pHnCBS1D plasmid (Figure 4-4B) which added the *csoS1D* gene into the pHnCB plasmid was shown to possess increased structural stability but decreased Rubisco activity α-carboxysomes, given that CsoS1D was deduced to gate the transport on the carboxysomes shell (Bonacci et al., 2012). This implies the possibility of improving the quality of synthetic carboxysomes by adding relevant components.

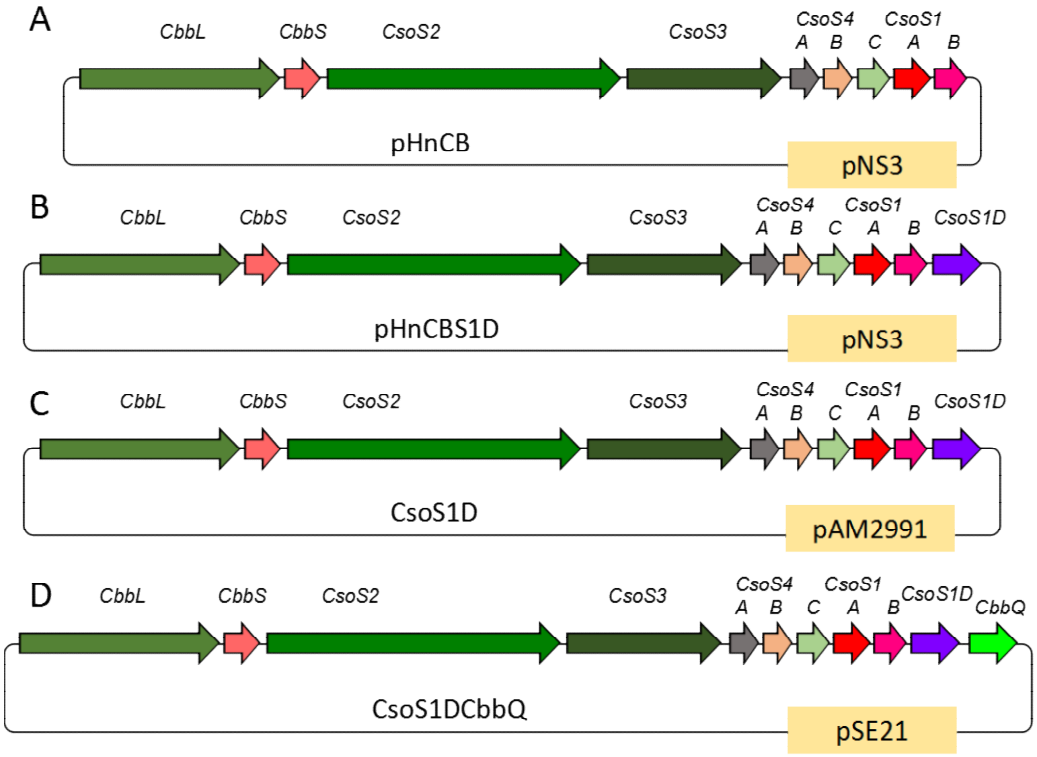

Figure 4-4. Synthetic operon of the *H. neapolitanus* α-carboxysome expressed in *E. coli* with different modification proteins. A:pHnCB and B:pHnCBS1D from previous research (Bonacci et al., 2012); C:CsoS1D and D: CsoS1DCbbQ from this project.

To enhance the structural stability and functionality of synthetic carboxysomes, I first inserted the *cso* operon and *CsoS1D* gene into a pAM2991 plasmid with a T7 promoter that can express proteins in *E.coli* by IPTG introduction (Figure 4-4C). After IPTG induction, this plasmid can induce the expression of carboxysome structures with stable Rubisco activity (Figure 4-8A). In addition, our cooperator group from Australia introduced CbbQ after the CsoS1D into the pSE21 (Grigor'Ev et al., 1990) plasmid to generate the plasmid named as CsoS1DCbbQ (Figure 4-4D).

Expression of CsoS1D and CsoS1DCbbQ in *E. coli* was induced by 50 μM IPTG which was the best condition to express carboxysomes by pHnCBS1D plasmid (Bonacci et al., 2012). The formation of synthetic α-carboxysomes was verified by thin-section electron microscopy. Expressed α-carboxysomes were densely packed in the cells containing CsoS1D (Figure 4-5A) and CsoS1DCbbQ (Figure 4-5B) plasmids.

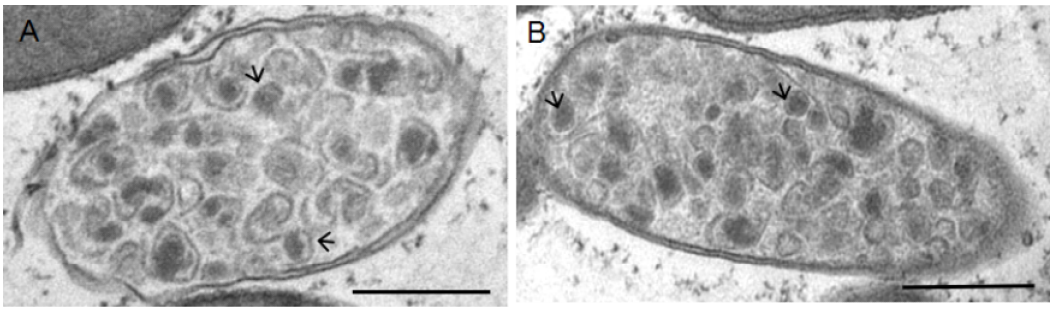

Figure 4-5. Electron microscopy of α-carboxysome *in vivo* (Scale bars: 500 nm).
A. Expression of CsoS1D at 50 μM IPTG.
B. Expression of CsoS1DCbbQ at 50 μM IPTG. Arrowheads in A and B indicated the carboxysomes *in vivo*.

4.2.2 Protein composition and structures of the purified synthetic α-carboxysomes

The recombinant α-carboxysomes were purified by sucrose density gradient centrifugation and were collected from the 40% fraction. SDS-PAGE of the CsoS1D and CsoS1DCbbQ isolated protein complexes showed similar composition as pHnCBS1D isolated samples (Figure 4-6A) (Bonacci et al., 2012), which revealed the presence of Rubisco subunits CbbLS, the major shell proteins CsoS1ABC and two types of CsoS2.

Both CsoS1D and CsoS1DCbbQ carboxysomes showed icosahedral shape under TEM (Figure 4-6B and 4-6C) and the diameters of the isolated CsoS1D and CsoS1DCbbQ carboxysomes were slightly larger than 100 nm, similar to that of pHnCBS1D (Bonacci et al., 2012). However, free Rubisco proteins (size ~ 10 nm and only exist in the isolated carboxysome samples) and broken shells \were observed in both isolated carboxysome samples (Figure 4-6B), which were also detected in pHnCBS1D isolated carboxysomes (Bonacci et al., 2012).

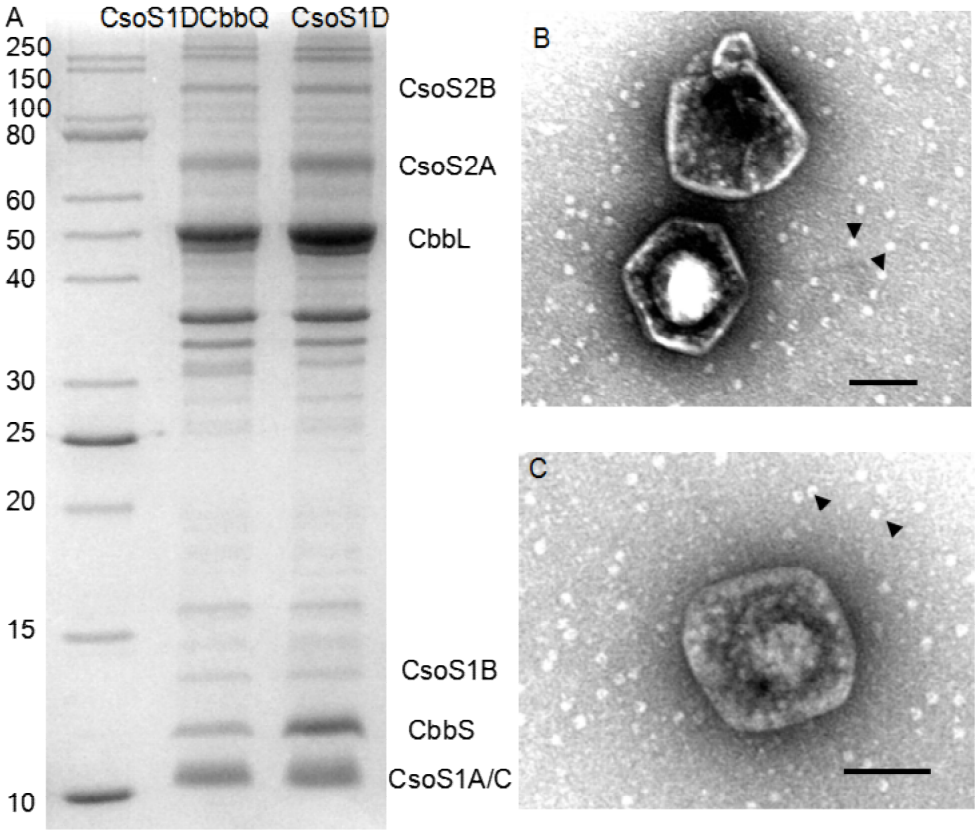

Figure 4-6. Characterization of purified α-carboxysomes *in vitro*.
A. SDS/PAGE gel of purified α-carboxysomes from CsoS1D, CsoS1DCbbQ strains.
B. Electron microscopy image of representative purified α-carboxysomes expressed using CsoS1D plasmid.
C. Electron microscopy image of representative purified α-carboxysomes expressed using CsoS1DCbbQ plasmid (Scale bars: 100 nm). Arrowheads in B and C indicated the free Rubisco proteins.

The purified samples were further analyzed by mass spectrometry, using yeast enolase (50 fmol μL^{-1}) as the control. The relative abundance of the proteins was normalized against the amount of the minor shell protein CsoS4A (Table 4-1). From the proteomic results, we found that all of the ten proteins from CsoS1D and 11 proteins from CsoS1DCbbQ were identified. The small shell protein CsoS1 is the most abundant, followed by Rubisco large subunits CbbL, and then Rubisco small subunits CbbS or shell associated protein CsoS2. CsoS4A/B is the minor proteins as

vertex were detected at lowest content, and the relative abundance of CsoS3 and CsoS1D were also minor.

Protein	CsoS1D Normalized amount (fmol)	CsoS1DCbbQ Normalized amount (fmol)
CsoS1	230.64±27.02	66.56±14.36
CbbL	125.59±56.34	208.23±79.51
CbbS	66.83±43.37	256.4±26.29
CsoS2	84.57±5.15	100.62±37.87
CsoS4B	1.63±0.19	1.29±0.08
CsoS1D	1.26±0.53	1.91±0.55
CsoS3	1.94±1.04	0.75±0.21
CsoS4A	1	1
CsoCbbQ	No CsoCbbQ gene	21.16±10.48

Table 4-1. Proteomic results of isolated α-carboxysome from CsoS1D and CsoS1DCbbQ.
The column of normalized amount displays the amount of each of the carboxysomal proteins detected in isolated α-carboxysome using mass spectroscopy, normalized against the amount of the CsoS4A.

We further studied the architecture of purified synthetic α-carboxysomes in solution using AFM. The synthetic CsoS1D carboxysomes (Figure 4-7A) were around 128.48 ± 21.51 nm high, while the synthetic CsoS1DCbbQ carboxysomes (Figure 4-7B) were around 110.30 ± 17.12 nm high, slightly smaller than CsoS1D carboxysomes.

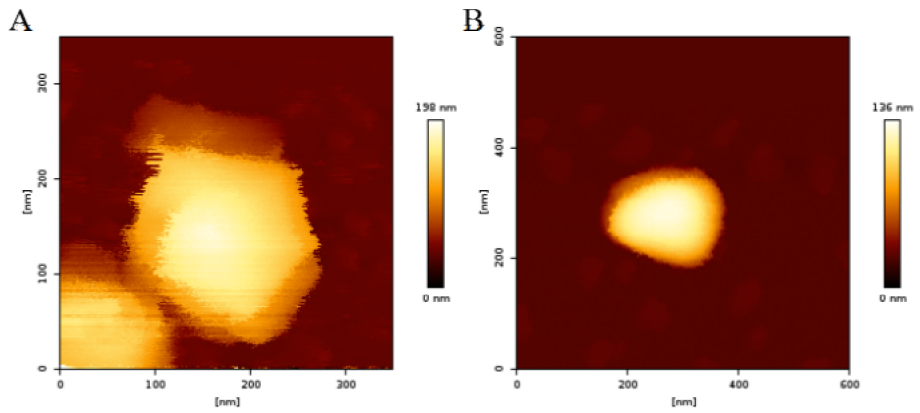

Figure 4-7. AFM image of intact purified α-carboxysome from A. CsoS1D and B. CsoS1DCbbQ.

4.2.3 Activity of synthetic α-carboxysomes

^{14}C radiometric Rubisco assays were performed to determine the carbon fixation activities of purified synthetic α-carboxysomes. $NaH^{14}CO_3$ can be dehydrated into $^{14}CO_2$ by CA enzyme, so by supplying a range of concentrations of D-ribulose 1,5-bisphosphate sodium salt hydrate (RuBP) as the substrate, the initial Rubisco fixation was measured by recording fixed ^{14}C in 3-PGA. Michaelis-Menten model was used to analyze the data. The CsoS1DCbbQ synthetic carboxysomes have a higher Rubisco activity with Vmax = 8.4 ± 1.6 μmol/min/mg (Figure 4-8 red) compared with CsoS1D Vmax = 2.9 ± 0.19 μmol/min/mg (Figure 4-8 black). Furthermore, the apparent Km for RuBP of CsoS1D carboxysomes is 0.23 ± 0.09 mM in contrast to that of CsoS1DCbbQ (Km = 0.96 ± 0.05 mM), which indicated higher RuBP affinity in CsoS1D carboxysomes than CsoS1DCbbQ carboxysomes.

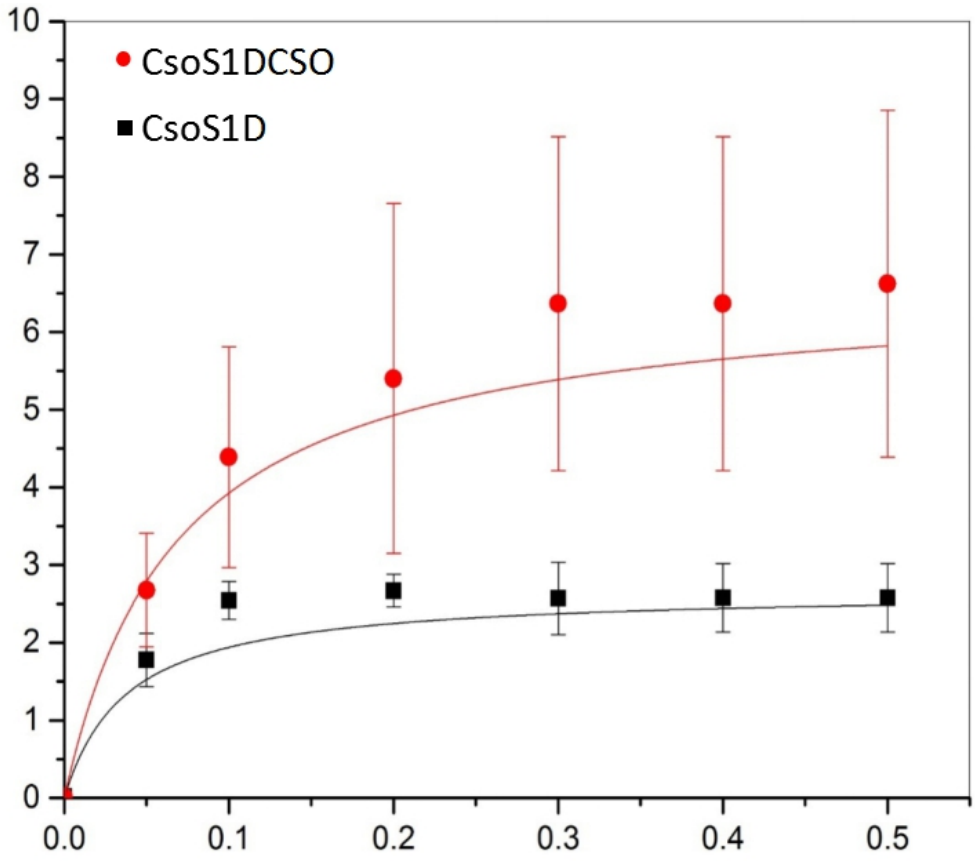

Figure 4-8. Carbon fixation activity of isolated α-carboxysome.
The Michaelis-Menten model to show the function of CsoS1D (black) and CsoS1DCbbQ (red) under a range of D-ribulose 1, 5-bisphosphateconcentration.

4.3 Discussion

Generating synthetic α-carboxysomes will help us to understand the essential proteins which functional carboxysomes required. Previous research on the heterologous production of α-carboxysomes was successful (Bonacci et al., 2012). Although CsoS1D as trimer improve the assembly of α-carboxysomes, the Rubisco activity of synthetic carboxysomes decreased compared with that of the carboxysomes expressed only *cso* operon (Bonacci et al., 2012).

The CbbQO complex from *H. neapolitanus* has been identified as Rubisco activase (Tsai et al., 2015; Sutter et al., 2015). CbbQ and CbbO can improve the free Rubisco activity in *E. coli* (Hayashi et al., 1997; Hayashi et al., 1999). In this study, we introduced CbbQ in α-carboxysome engineering. By the combination of molecular genetics, SDS-PAGE, Rubisco assays, proteomic analysis, electron and atomic force microscopy, we rebuilt the construct by adding CbbQ and produced heterologously α-carboxysomes in *E. coli*. The structure and activity of synthetic α-carboxysomes were characterized and validated. Our data showed the incorporation of CbbQ into the synthetic carboxysome. No detectable structural change of synthetic carboxysomes was observed. More importantly, there is a notable increase in Rubisco activity with the addition of CbbQ, which demonstrates its function as the Rubisco activase. However, the effects of addition of CbbO and other factors on improving carboxysome structure, assembly and activity require further investigation.

4.4 References

Andralojc, P. J., P. J. Madgwick, Y. Tao, A. Keys, J. L. Ward, M. H. Beale, J. E. Loveland, P. J. Jackson, A. C. Willis, S. Gutteridge, and M. A. Parry. 2012. "2-Carboxy-D-arabinitol 1-phosphate (CA1P) phosphatase: evidence for a wider role in plant Rubisco regulation", *Biochem J*, 442: 733-42.

Badger, M. R., and E. J. Bek. 2008. "Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO_2 acquisition by the CBB cycle", *J Exp Bot*, 59: 1525-41.

Badger, M. R., and G. D. Price. 2003. "CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution", *J Exp Bot,* 54: 609-22.

Baker, S. H., S. C. Lorbach, M. Rodriguez-Buey, D. S. Williams, H. C. Aldrich, and J. M. Shively. 1999. "The correlation of the gene csoS2 of the carboxysome operon

with two polypeptides of the carboxysome in *Thiobacillus neapolitanus*", *Arch Microbiol*, 172: 233-9.

Baumgart, M., I. Huber, I. Abdollahzadeh, T. Gensch and J. Frunzke. 2017. "Heterologous expression of the *Halothiobacillus neapolitanus* carboxysomal gene cluster in *Corynebacterium glutamicum*." *J Biotechnol* 258: 126-135.

Bhat, J. Y., G. Thieulin-Pardo, F. U. Hartl, and M. Hayer-Hartl. 2017. "Rubisco Activases: AAA^+ Chaperones Adapted to Enzyme Repair", *Front Mol Biosci,* 4: 20.

Bonacci, W., P. K. Teng, B. Afonso, H. Niederholtmeyer, P. Grob, P. A. Silver, and D. F. Savage. 2012. "Modularity of a carbon-fixing protein organelle", *Proc Natl Acad Sci U S A*, 109: 478-83.

Cai, F., Z. Dou, S. L. Bernstein, R. Leverenz, E. B. Williams, S. Heinhorst, J. Shively, G. C. Cannon, and C. A. Kerfeld. 2015. "Advances in Understanding Carboxysome Assembly in *Prochlorococcus* and *Synechococcus* Implicate CsoS2 as a Critical Component", *Life*, 5: 1141-71.

Cai, F., B. B. Menon, G. C. Cannon, K. J. Curry, J. M. Shively, and S. Heinhorst. 2009. "The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO_2 leakage barrier", *PLoS One*, 4: e7521.

Cannon, G. C., S. H. Baker, F. Soyer, D. R. Johnson, C. E. Bradburne, J. L. Mehlman, P. S. Davies, Q. L. Jiang, S. Heinhorst, and J. M. Shively. 2003. "Organization of carboxysome genes in the thiobacilli", *Curr Microbiol*, 46: 115-9.

Cannon, G. C., C. E. Bradburne, H. C. Aldrich, S. H. Baker, S. Heinhorst, and J. M. Shively. 2001. "Microcompartments in prokaryotes: carboxysomes and related polyhedra", *Appl Environ Microbiol,* 67: 5351-61.

Chaijarasphong, T., R. J. Nichols, K. E. Kortright, C. F. Nixon, P. K. Teng, L. M. Oltrogge, and D. F. Savage. 2016. "Programmed Ribosomal Frameshifting Mediates Expression of the alpha-Carboxysome", *J Mol Biol*, 428: 153-64.

Dou, Z., S. Heinhorst, E. B. Williams, C. D. Murin, J. M. Shively, and G. C. Cannon. 2008. "CO_2 fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO_2", *J Biol Chem*, 283: 10377-84.

English, R. S., S. C. Lorbach, X. Qin, and J. M. Shively. 1994. "Isolation and characterization of a carboxysome shell gene from *Thiobacillus neapolitanus*", *Mol Microbiol*, 12: 647-54.

Gonzales, Arlene D, Yooli K Light, Zhaoduo Zhang, Tahera Iqbal, Todd W Lane, and Anthony Martino. 2005. "Proteomic analysis of the CO_2-concentrating mechanism in the open-ocean cyanobacterium *Synechococcus* WH8102", *Canadian journal of botany,* 83: 735-45.

Grigor'Ev, A. E., T. S. Belova, M. I. Naĭdenova and V. N. Danilenko. 1990. "Plasmid pSE21 of Streptomyces erythraeus strains." *Antibiot Khimioter*. Apr;35(4):6-10.

Hayashi, Nobuhiro R, Hiroyuki Arai, Tohru Kodama, and Yasuo Igarashi. 1997. "The Novel Genes, cbbQ and cbbO, Located Downstream from the RubisCO Genes of Pseudomonas hydrogenothermophila, Affect the Conformational States and Activity of RubisCO", *Biochemical and biophysical research communications,* 241: 565-69.

Hayashi N. R., Arai H., Kodama T. & Igarashi Y. 1999. "The cbbQ genes, located downstream of the form I and form II RubisCO genes, affect the activity of both RubisCOs". *Biochem. Biophys. Res. Commun*. 265, 177–183.

Heinhorst, S., E. B. Williams, F. Cai, C. D. Murin, J. M. Shively, and G. C. Cannon. 2006. "Characterization of the carboxysomal carbonic anhydrase CsoSCA from *Halothiobacillus neapolitanus*", *J Bacteriol*, 188: 8087-94.

Holthuijzen, Yolande A, Jan FL van Breemen, J Gijs Kuenen, and Wil N Konings. 1986. "Protein composition of the carboxysomes of *Thiobacillus neapolitanus*", *Archives of Microbiology*, 144: 398-404.

Iancu, C. V., H. J. Ding, D. M. Morris, D. P. Dias, A. D. Gonzales, A. Martino, and G. J. Jensen. 2007. "The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography", *J Mol Biol*, 372: 764-73.

Iancu, C. V., D. M. Morris, Z. Dou, S. Heinhorst, G. C. Cannon, and G. J. Jensen. 2010. "Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells", *J Mol Biol*, 396: 105-17.

Jordan, D. B., and R. Chollet. 1983. "Inhibition of ribulose bisphosphate carboxylase by substrate ribulose 1,5-bisphosphate", *J Biol Chem*, 258: 13752-8.

Kelly, Donovan P, and Ann P Wood. 2000. "Reclassification of some species of Thiobacillus to the newly designated genera Acidithiobacillus gen. nov., Halothiobacillus gen. nov. and Thermithiobacillus gen. nov", *International Journal of Systematic and Evolutionary Microbiology*, 50: 511-16.

Kerfeld, C. A., and M. R. Melnicki. 2016. "Assembly, function and evolution of cyanobacterial carboxysomes", *Curr Opin Plant Biol*, 31: 66-75.

Klein, M. G., P. Zwart, S. C. Bagby, F. Cai, S. W. Chisholm, S. Heinhorst, G. C. Cannon, and C. A. Kerfeld. 2009. "Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport", *J Mol Biol*, 392: 319-33.

Lorimer, G. H., M. R. Badger, and T. J. Andrews. 1976. "The activation of ribulose-1,5-bisphosphate carboxylase by carbon dioxide and magnesium ions. Equilibria, kinetics, a suggested mechanism, and physiological implications", *Biochemistry*, 15: 529-36.

Menon, B. B., Z. Dou, S. Heinhorst, J. M. Shively, and G. C. Cannon. 2008. "*Halothiobacillus neapolitanus* carboxysomes sequester heterologous and chimeric RubisCO species", *PLoS One*, 3: e3570.

Mueller-Cajar, O., M. Stotz, P. Wendler, F. U. Hartl, A. Bracher, and M. Hayer-Hartl. 2011. "Structure and function of the AAA^+ protein CbbX, a red-type Rubisco activase", *Nature*, 479: 194-9.

Parry, M. A., A. J. Keys, P. J. Madgwick, A. E. Carmo-Silva, and P. J. Andralojc. 2008. "Rubisco regulation: a role for inhibitors", *J Exp Bot*, 59: 1569-80.

Pena, K. L., S. E. Castel, C. de Araujo, G. S. Espie and M. S. Kimber. 2010. "Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM." *Proc Natl Acad Sci U S A* 107(6): 2455-2460.

Price, G. D., and M. R. Badger. 1989. "Expression of Human Carbonic Anhydrase in the Cyanobacterium *Synechococcus* PCC7942 Creates a High CO_2-Requiring Phenotype : Evidence for a Central Role for Carboxysomes in the CO(2) Concentrating Mechanism", *Plant Physiol*, 91: 505-13.

Price, G. D., S. M. Howitt, K. Harrison, and M. R. Badger. 1993. "Analysis of a genomic DNA region from the cyanobacterium *Synechococcus* sp. strain PCC7942 involved in carboxysome assembly and function", *J Bacteriol*, 175: 2871-9.

Rae, B. D., B. M. Long, M. R. Badger, and G. D. Price. 2013. "Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria", *Microbiol Mol Biol Rev*, 77: 357-79.

Roberts, E. W., F. Cai, C. A. Kerfeld, G. C. Cannon, and S. Heinhorst. 2012. "Isolation and characterization of the *Prochlorococcus* carboxysome reveal the presence of the novel shell protein CsoS1D", *J Bacteriol*, 194: 787-95.

Santelli, Eugenio, Laurie A Bankston, Stephen H Leppla, and Robert C Liddington. 2004. "Crystal structure of a complex between anthrax toxin and its host cell receptor", *Nature*, 430: 905.

Sawaya, M. R., G. C. Cannon, S. Heinhorst, S. Tanaka, E. B. Williams, T. O. Yeates, and C. A. Kerfeld. 2006. "The structure of beta-carbonic anhydrase from the carboxysomal shell reveals a distinct subclass with one active site for the price of two", *J Biol Chem*, 281: 7546-55.

Schmid, M. F., A. M. Paredes, H. A. Khant, F. Soyer, H. C. Aldrich, W. Chiu, and J. M. Shively. 2006. "Structure of *Halothiobacillus neapolitanus* carboxysomes by cryo-electron tomography", *J Mol Biol*, 364: 526-35.

So, A. K., G. S. Espie, E. B. Williams, J. M. Shively, S. Heinhorst, and G. C. Cannon. 2004. "A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell", *J Bacteriol*, 186: 623-30.

Sutter, M., E. W. Roberts, R. C. Gonzalez, C. Bates, S. Dawoud, K. Landry, G. C. Cannon, S. Heinhorst, and C. A. Kerfeld. 2015. "Structural Characterization of a Newly Identified Component of alpha-Carboxysomes: The AAA^+ Domain Protein CsoCbbQ", *Sci Rep*, 5: 16243.

Tanaka, S., C. A. Kerfeld, M. R. Sawaya, F. Cai, S. Heinhorst, G. C. Cannon, and T. O. Yeates. 2008. "Atomic-level models of the bacterial carboxysome shell", *Science*, 319: 1083-6.

Teru O. and A.J.Wilkinson. 2001. AAA^+ superfamily ATPases: common structure – diverse function. *Genes to Cells*, 6(7), pp.575-597.

Tsai, Y. C., M. C. Lapina, S. Bhushan, and O. Mueller-Cajar. 2015. "Identification and characterization of multiple rubisco activases in chemoautotrophic bacteria", *Nat Commun*, 6: 8883.

Tsai, Y., M. R. Sawaya, G. C. Cannon, F. Cai, E. B. Williams, S. Heinhorst, C. A. Kerfeld, and T. O. Yeates. 2007. "Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome", *PLoS Biol*, 5: e144.

Wang, L., T. Yamano, M. Kajikawa, M. Hirono, and H. Fukuzawa. 2014. "Isolation and characterization of novel high-CO_2-requiring mutants of *Chlamydomonas reinhardtii*", *Photosynth Res*, 121: 175-84.

Wong, Keith S, and Walid A Houry. 2012. "Novel structural and functional insights into the MoxR family of AAA^+ ATPases", *Journal of structural biology*, 179: 211-21.

Xiong, J. P., T. Stehle, R. Zhang, A. Joachimiak, M. Frech, S. L. Goodman, and M. A. Arnaout. 2002. "Crystal structure of the extracellular segment of integrin alpha V beta3 in complex with an Arg-Gly-Asp ligand", *Science*, 296: 151-5.

Zarzycki, J., S. D. Axen, J. N. Kinney, and C. A. Kerfeld. 2013. "Cyanobacterial-based approaches to improving photosynthesis in plants", *J Exp Bot*, 64: 787-98.

Chapter 5

Engineering α-carboxysome shells in *E. coli*: a new nanobioreactors

5.1 Introduction

The protein shell of BMCs, structurally resembling virus capsids, is made of hundreds of multiple protein paralogs forming hexagons and pentagons and acts as a physical barrier that controls the passage of substrates and products of enzymatic reactions. These specific structures allow BMCs to reduce toxicity and increase the activating of entire metabolic pathway. All of the shells of BMCs share an architectural similarity as I described before (Chapter 1-1.1, Figure 1-2)(Kerfeld et al., 2015; Yeates et al., 2013).

There are mainly three types of proteins that build up BMC shells, the hexamers as shell proteins, the pentamers located at the vertices of the shell and the trimers or pseudohexamers deduced to act as the barriers for controlling metabolite transport (Yeates et al., 2013). Depending on the native structure of BMC, we can enhance the catalytic metabolism by imitating the structure or even redesign them to receive new synthetic purposes (Jordan et al., 2016).

The engineering of synthetic BMC shells has drawn particular attention recently for bioengineering of nanoreactors (Choudhary, 2012; Parsons, 2010; Parsons, 2008; Sargent, 2013). The shell of β-carboxysomes from *Halothece sp.* PCC 7418 were synthesized in *E. coli* by using two BMC-H proteins CcmK1 and CcmK2, one BMC-P protein CcmL and one BMC-T protein CcmO (Cai et al., 2016). These shell proteins form 20 nm shell structures containing a single protein layer, notably smaller than that of native carboxysomes (~150 nm) (Cai et al., 2016). Furthermore, fluorescent

proteins can be recruited by using the encapsulation peptide of CcmN or the N-term of CcmM (Cai et al., 2016).

The synthetic EUT and PDU shells have also been produced and can encapsulate fluorescent proteins or other enzymes in *E. coli*. Hexameric EutS protein alone or EutSMNLK proteins together can self-assemble to form polyhedral shells, and EutQ serves as an assembly hub to control the number of synthetic shells (Choudhary et al., 2012; Held et al., 2016). The shell proteins of PDU expressed in *E. coli* independently can form different aggregated structures. The combination of PduABJKMNTU can create a minimal PDU shell (Parson et al., 2010; Sinha et al., 2012; Wagner et al., 201; Parsons et al., 2008). The β-galactosidase, esterase Est5, and glycerol dehydrogenase have been fused with PduP-EP and captured inside of synthetic PDU shell in catalytically active form (Hanna et al., 2017). Furthermore, although the function of the microcompartments from *Haliangium ochraceum* hasn't been characterized, the shell proteins can also produce small shells in high yield to encapsulate new proteins by EP tags (Lassila et al., 2014; Sutter et al., 2017).

There are also other cage-like architectures built up based on self-assembling proteins, some of which have shown to encapsulate active enzymes (Fiedler et al., 2010; Comellas et al., 2007; Worsdorfer et al., 2011; Inoue et al., 2008; Snijder et al., 2016). The lumazine synthases from *Aquifex aeolicus*, which catalyze the penultimate step of riboflavin biosynthesis, are also used to build protein cages (Frey et al., 2016), and this protein cage can encapsulate carboxysomal enzymes Rubisco and CA by electrostatic tagging system to form mimic carboxysomes. Although these mimic

carboxysomes have no specific permeability to HCO_3^- or CO_2, this research still showed the principles for constructing artificial nanocages (Frey et al., 2016).

Using computational design, 60-subunit icosahedral nanocages (25 nm) are generated in *E. coli* by the self-assembled trimeric protein (Hsia et al., 2016). GFP can be fused to the shell, and designed protein pentamers can be placed on it to modulate the size of the cage. Furthermore, the combination of different ratio of dimeric, trimeric and pentameric proteins can result in the construction of 120-subunit nanocages. Three distinct icosahedral architectures are generated and they are large enough to capture nucleic acids or proteins inside (Bale et al., 2016). Furthermore, the oxygen-tolerant [NiFE]- hydrogenase has been encapsulated within the bacteriophage P22 capsid and shown to be functional in hydrogen production (Jordan et al., 2016).

To study the synthetic α-carboxysome shells but not the native α-carboxysome shells in *H. neapolitanus* (Menon et al., 2008) will help us to understand whether the self-assembly of the carboxysome shells requires any other help. And it is much easier to modify the components and the structure of the shells or introduce different cargos inside the shells in *E.coli* than in *H. neapolitanus*. In this chapter, we generated a construct to express synthetic α-carboxysomes without Rubisco, allowing the production of empty α-carboxysome shells in *E. coli*. We isolated the α-carboxysome shells and characterized their protein composition, structure, and stability. Building synthetic carboxysomes shell is vital for understanding the assembly of BMC shell and its potential to recruit nanoreactors for agricultural, industrial or medical use.

5.2 Results

5.2.1 The synthetic operon for expressing α-carboxysome shell proteins

The α-carboxysomes major proteins were all in one major *cso* operon. Deletion of Rubisco in *H. neapolitanus* impact little on the structure and the size of the shell (Menon et al., 2008). Our CsoS1D plasmid built in Chapter 4 contains genes encoding all the shell proteins, Rubisco, and CA. To acquire a plasmid which can synthesize the empty shell, we knocked out the *cbbL* and *cbbS* genes from this plasmid (Figure 5-1). First, the primers contain 39 bp overlap sequences before *CbbL* and after *CbbS* genes were used to amplify the *aadA* gene (spectinomycin-resistance). Then, with the help of λ-Red recombinase plasmid pKD46 (Baba et al.,2006; Datsenko et al, 2000), *aadA* gene was introduced into the CsoS1D plasmid to replace the *cbbL* and *cbbS* genes. Finally, the pCP20 Flp helper plasmid (Cherepanov et al., 1995) were introduced to delete the aadA gene by Flp-FRT (recombinase flippase - flippase recognition target) recombination (Zhu et al., 1995). The PCR results from *CbbL: CbbS* primers showed there are no *CbbL: CbbS* genes left (Figure 5-2A right), and *CsoS3: CsoS4* PCR has bands means other parts of the plasmid are intact (Figure 5-2A left). So by the two-step recombination, we obtained the α-carboxysome shell expression plasmid consisting of genes encoding major shell proteins CsoS1A/B/C, minor shell proteins CsoS4A/B, shell-associated proteins CsoS2 and CsoS3 (Figure 5-1). The α-carboxysome shell was produced using the same protocol for heterologous expression of α-carboxysomes described in Chapter 4: 50 μM IPTG, 18ºC treatment for overnight.

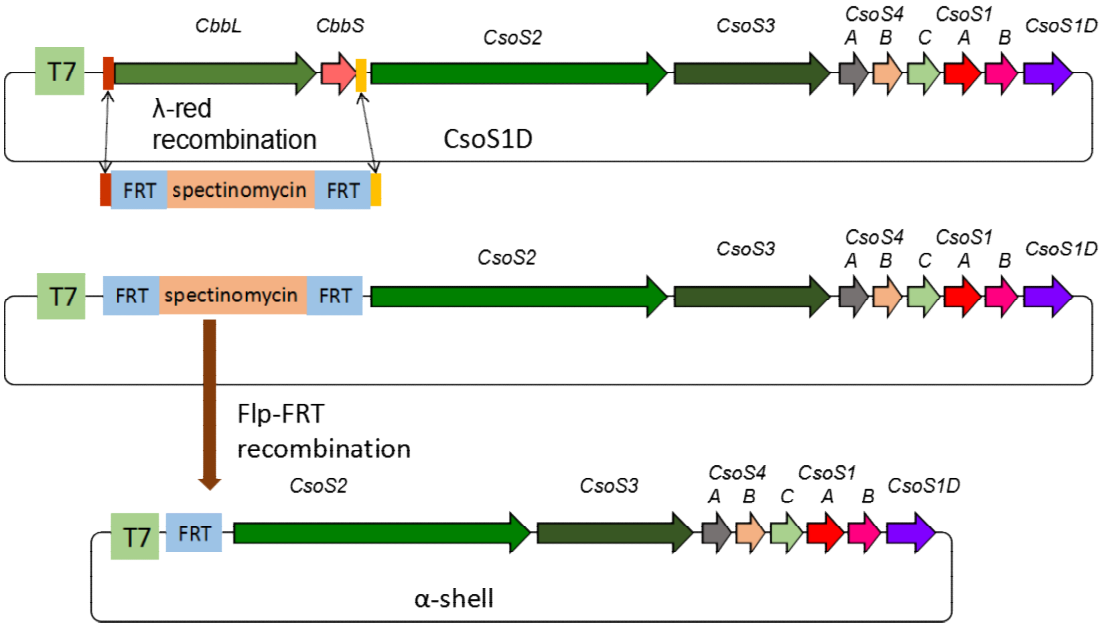

Figure 5-1. The synthetic operon for expression of the *H. neapolitanus* shell proteins by two-step recombination to knock out Rubisco from CsoS1D plasmid and obtain shell plasmid.
First use spectinomycin to replace *cbbL:cbbS* by using BW25113 strain which carries λ-Red recombinase plasmid pKD46. And then use the FRT site to cut spectinomycin to get the clean α-carboxysome shell plasmid.

Shell-expressing cells were analyzed using thin-section EM. We observed shell-like structures inside the *E. coli* cell (Figure 5-2B). They are not like synthetic α-carboxysomes *in vivo* with high internal protein density but appear as curly structures or empty carboxysome shell with broken structures.

5.2.2 Purification of α-carboxysome shells

The α-carboxysome shells in *H. neapolitanus* were produced by replacing Rubisco with kanamycin, and the empty shells can be isolated using the same protocol for isolating native α-carboxysomes in *H. neapolitanus* (Menon et al., 2008). Thus, we used the method for purifying synthetic α-carboxysomes to purify synthetic shells and

harvest the 20% sucrose fractions, because the empty shells should be lighter than the intact carboxysomes. The SDS-PAGE results showed strong expression of CsoS1A/B/C and CsoS2A/B proteins, and the absence of CbbL and CbbS protein bands were observed when compared with CsoS1D results (Figure 5-2C). CsoS3, CsoS4A/B, and CsoS1D were not observed in SDS-PAGE, likely due to the low expression levels. These results indicate that the shell proteins can self-assemble together.

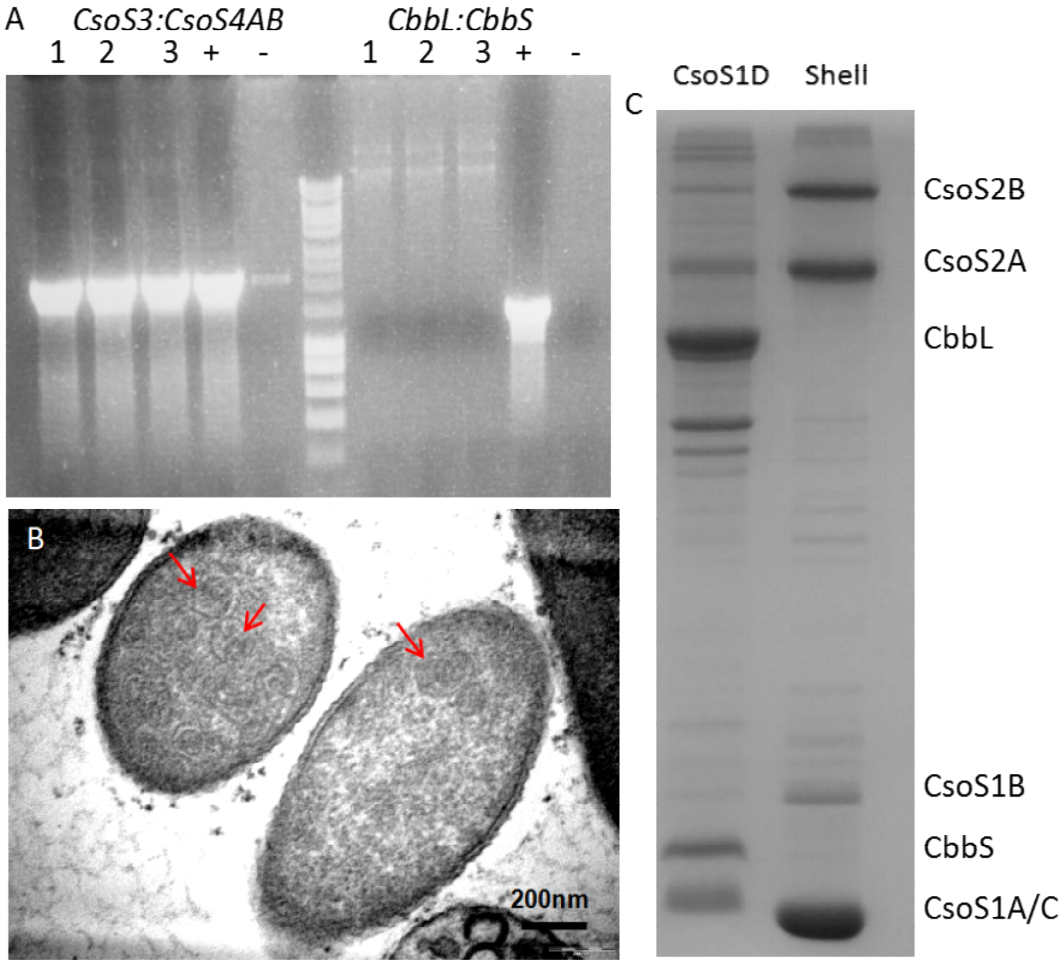

Figure 5-2. Expression of α-carboxysome shell.
A. PCR results of the knock out Rubisco from CsoS1D strain to get the α-shell plasmid. Primers used to detect *CsoS3:CsoS4AB* and *CbbL:CbbS* were in the Table 2-1. "+" represents positive control, "-" represents negative control
B. Thin-section EM image showing induced *E. coli* cells, arrowheads showed foreigner structures. Scar bar=200 nm
C. SDS-PAGE of purified shell from 20% fraction of sucrose gradient, CsoS1A/B/C and CsoS2 can be detected. Purified CsoS1D from 40% fraction as control showed CbbL and CbbS proteins.

The isolated α-carboxysome shells obtained from 20% sucrose fraction were analyzed by mass spectroscopy, and the results showed that all the shell proteins were expressed. The highest abundant proteins are CsoS1A/B/C and followed by CsoS2 protein; CsoS4A/B, CsoS1D and CsoS3 proteins possess low abundance (Table 5-1). The proteomic analysis demonstrated that the ratio of the component and

the constitution of the empty shell is similar to the synthetic α-carboxysomes (Table 5-1). No Rubisco proteins were detected, suggesting that Rubisco is not essential for α-carboxysome shell assembly. This is distinct from the biogenesis pathway of β-carboxysomes (Cameron et al., 2013).

Protein	Normalized amount(fmol)
CsoS1	101.76 ± 9.5
CsoS2	48.62 ± 8.26
CsoS4B	2.64 ± 0.03
CsoS1D	0.16 ± 0.05
CsoS3	0.16 ± 0.08
CsoS4A	1

Table 5-1. Proteomic results of the isolated shell.
The column of the normalized amount displays the amount of each of the carboxysomal proteins detected in the isolated shell using mass spectroscopy, normalized against the amount of CsoS4A.

5.2.3 Characterization of purified α-carboxysome shells

The 20% fractions of sucrose gradient were examined by TEM. In the 20% sucrose fraction, icosahedral structures with edges and vertexes were found (Figure 5-3A). Compared with synthetic α-carboxysomes, this structure has a low protein density within the compartment. Furthermore, the size of the particles is around 60~80 nm (Figure 5-3B), smaller than synthetic and native α-carboxysomes (Chapter 4; Bonacci et al., 2012; Faulkner et al., 2017). These TEM results showed the empty α-carboxysome shell could be heterologously assembled successfully. Compared

with 20% sucrose fraction, the 30% and 40% sucrose fractions have lower protein density and fewer carboxysome-like structures were observed in EM, by contrast, synthetic α-carboxysomes were mainly distributed in the 40% sucrose fraction (Chapter 4), due to the low protein density of synthetic shells.

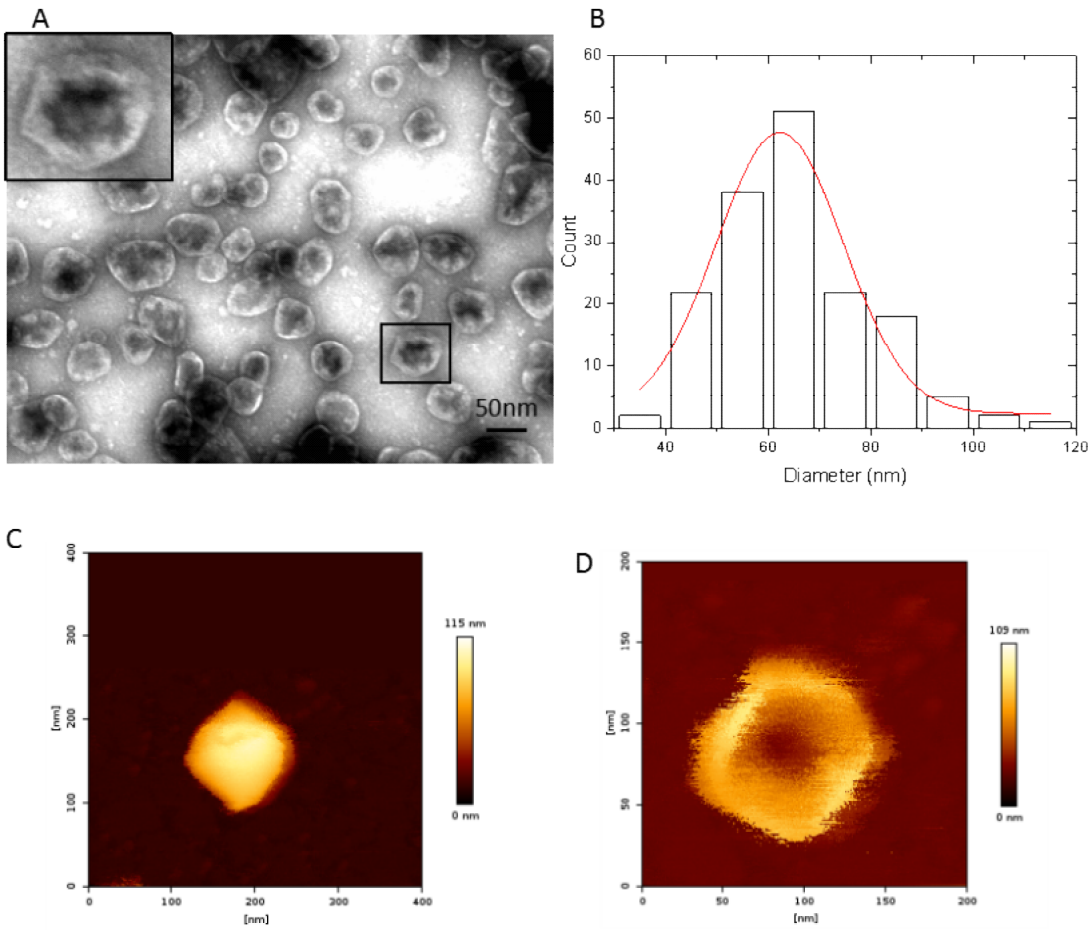

Figure 5-3. Characterization of the purified α-carboxysome shells.
A. Negatively stained electron microscopy images of the isolated α-carboxysome shells. (*Inset*) Zoom in of the shell. Scar bar = 50 nm.
B. The distribution of the diameter of isolated α-carboxysome shell depends on 150 particles.
C. AFM image of intact purified α-carboxysome shell and D. after breakage to show the empty inside of the shell.

We further characterized the shell structure by AFM. The empty shell showed an average height of 50~80 nm (Figure 5-3C). We also found that the shells can be

easily broken during AFM imaging, even with gentle scanning force less than 100 pN. After the shell breakage, internal hollow of the shell was visible (Figure 5-3D). The softness of the shell might be ascribed to the lack of Rubisco packing.

5.3 Discussion

The synthetic shells enable systematic study on the structural and functional roles of shell proteins. We found that empty α-carboxysome shells can be expressed and formed in *E. coli* without Rubisco, indicating the independence of α-carboxysome shell formation on the presence of Rubisco core. Furthermore, the lack of the core within the shell might be the reason for the reduced size of the empty shell in contrast to those of synthetic and native α-carboxysomes (Bonacci et al., 2012; Faulkner et al., 2017).

The production of synthetic α-carboxysome shells provides promise for design and engineering new bioreactors by recruiting new enzymes. The carboxysomes shell create a microenvironment with a high CO2 concentration around the cargos, which may facilitate the activity of O_2-sensitive enzymes. For example, hydrogenases are sensitive to O_2 and can transform H_2O into H_2. By linking scaffold proteins with two major subunits of hydrogenase, HyaA and HyaB, the hydrogenase could be captured and kept functional inside bacteriophage P22 capsid shells (Jordan et al., 2016). It was shown that the C-terminus of CsoS2 serves as the linker to target Rubisco to the shell (Cai et al., 2015). Thus, fusing hydrogenase with the C-terminus of CsoS2 may

allow the generation of new nanocages encapsulating hydrogenases to underpin hydrogen production, which has potential industrial applications.

5.4 References

Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. 2006. "Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection". *Molecular Systems Biology*. 2: 2006.0008.

Bale, J. B., S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King and D. Baker. 2016. "Accurate design of megadalton-scale two-component icosahedral protein complexes." *Science* 353(6297): 389-394.

Bonacci, W., P. K. Teng, B. Afonso, H. Niederholtmeyer, P. Grob, P. A. Silver and D. F. Savage. 2012. "Modularity of a carbon-fixing protein organelle." *Proc Natl Acad Sci U S A* 109(2): 478-483.

Cai, F., S. L. Bernstein, S. C. Wilson and C. A. Kerfeld. 2016. "Production and Characterization of Synthetic Carboxysome Shells with Incorporated Luminal Proteins." *Plant Physiol* 170(3): 1868-1877.

Cai, F., Z. Dou, S. L. Bernstein, R. Leverenz, E. B. Williams, S. Heinhorst, J. Shively, G. C. Cannon and C. A. Kerfeld. 2015. "Advances in Understanding Carboxysome Assembly in *Prochlorococcus* and *Synechococcus* Implicate CsoS2 as a Critical Component." *Life* 5(2): 1141-1171.

Cameron, J. C., S. C.Wilson, S. L. Bernstein, and C. A. Kerfeld. 2013. "Biogenesis of a bacterial organelle: the carboxysome assembly pathway." *Cell* 155(5): 1131-1140.

Cherepanov, P. P. and W. Wackernagel. 1995. "Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant." *Gene* 158(1): 9-14.

Choudhary, S., M. B. Quin, M. A. Sanders, E. T. Johnson and C. Schmidt-Dannert. 2012. "Engineered protein nano-compartments for targeted enzyme localization." *PLoS One* 7(3): e33342.

Comellas-Aragones, M., H. Engelkamp, V. I. Claessen, N. A. Sommerdijk, A. E. Rowan, P. C. Christianen, J. C. Maan, B. J. Verduin, J. J. Cornelissen and R. J. Nolte. 2007. "A virus-based single-enzyme nanoreactor." *Nat Nanotechnol* 2(10): 635-639.

Datsenko, K. A. and B. L. Wanner. 2000. "One-step inactivation of chromosomal

genes in *Escherichia coli* K-12 using PCR products." *Proceedings of the National Academy of Sciences* 97(12): 6640-6645.

Faulkner, M., J. Rodriguez-Ramos, G. F. Dykes, S. V. Owen, S. Casella, D. M. Simpson, R. J. Beynon and L. N. Liu. 2017. "Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes." *Nanoscale* 9(30): 10662-10673.

Fiedler, J. D., S. D. Brown, J. L. Lau and M. G. Finn. 2010. "RNA-directed packaging of enzymes within virus-like particles." *Angew Chem Int Ed Engl* 49(50): 9648-9651.

Frey, R., S. Mantri, M. Rocca and D. Hilvert. 2016. "Bottom-up Construction of a Primordial Carboxysome Mimic." *J Am Chem Soc* 138(32): 10072-10075.

Held, M., A. Kolb, S. Perdue, S. Y. Hsu, S. E. Bloch, M. B. Quin and C. Schmidt-Dannert. 2016. "Engineering formation of multiple recombinant Eut protein nanocompartments in *E. coli*." *Sci Rep* 6: 24359.

Hsia, Y., J. B. Bale, S. Gonen, D. Shi, W. Sheffler, K. K. Fong, U. Nattermann, C. Xu, P. S. Huang, R. Ravichandran, S. Yi, T. N. Davis, T. Gonen, N. P. King and D. Baker. 2016. "Design of a hyperstable 60-subunit protein dodecahedron." *Nature* 535(7610): 136-139.

Inoue, T., M. A. Kawano, R. U. Takahashi, H. Tsukamoto, T. Enomoto, T. Imai, K. Kataoka and H. Handa. 2008. "Engineering of SV40-based nano-capsules for delivery of heterologous proteins as fusions with the minor capsid proteins VP2/3." *J Biotechnol* 134(1-2): 181-192.

Jordan, P. C., D. P. Patterson, K. N. Saboda, E. J. Edwards, H. M. Miettinen, G. Basu, M. C. Thielges and T. Douglas. 2016. "Self-assembling biomolecular catalysts for hydrogen production." *Nat Chem* 8(2): 179-185.

Kerfeld, C. A. and O. Erbilgin. 2015. "Bacterial microcompartments and the modular construction of microbial metabolism." *Trends Microbiol* 23(1): 22-34.

Lassila, J. K., S. L. Bernstein, J. N. Kinney, S. D. Axen and C. A. Kerfeld. 2014. "Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function." *J Mol Biol* 426(11): 2217-2228.

Menon, B. B., Z. Dou, S. Heinhorst, J. M. Shively and G. C. Cannon. 2008. "*Halothiobacillus neapolitanus* carboxysomes sequester heterologous and chimeric RubisCO species." *PLoS One* 3(10): e3570.

Parsons, J. B., S. D. Dinesh, E. Deery, H. K. Leech, A. A. Brindley, D. Heldt, S. Frank, C. M. Smales, H. Lünsdorf and A. Rambach. 2008. "Biochemical and structural insights into bacterial organelle form and biogenesis." *Journal of biological chemistry* 283(21): 14366-14375.

Parsons, J. B., S. Frank, D. Bhella, M. Liang, M. B. Prentice, D. P. Mulvihill and M. J.

Warren. 2010. "Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement." *Mol Cell* 38(2): 305-315.

Sargent, F., F. A. Davidson, C. L. Kelly, R. Binny, N. Christodoulides, D. Gibson, E. Johansson, K. Kozyrska, L. L. Lado, J. Maccallum, R. Montague, B. Ortmann, R. Owen, S. J. Coulthurst, L. Dupuy, A. R. Prescott and T. Palmer. 2013. "A synthetic system for expression of components of a bacterial microcompartment." *Microbiology* 159(Pt 11): 2427-2436.

Sinha, S., S. Cheng, C. Fan and T. A. Bobik. 2012. "The PduM protein is a structural component of the microcompartments involved in coenzyme B12-dependent 1, 2-propanediol degradation by *Salmonella enterica*." *Journal of bacteriology* 194(8): 1912-1918.

Snijder, J., O. Kononova, I. M. Barbu, C. Uetrecht, W. F. Rurup, R. J. Burnley, M. S. Koay, J. J. Cornelissen, W. H. Roos, V. Barsegov, G. J. Wuite and A. J. Heck. 2016. "Assembly and Mechanical Properties of the Cargo-Free and Cargo-Loaded Bacterial Nanocompartment Encapsulin." *Biomacromolecules* 17(8): 2522-2529.

Sutter, M., B. Greber, C. Aussignargues and C. A. Kerfeld. 2017. "Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell." *Science* 356(6344): 1293-1297.

Wagner, H. J., C. C. Capitain, K. Richter, M. Nessling and J. Mampel. 2017. "Engineering bacterial microcompartments with heterologous enzyme cargos." *Engineering in Life Sciences* 17(1): 36-46.

Worsdorfer, B., K. J. Woycechowsky and D. Hilvert. 2011. "Directed evolution of a protein container." *Science* 331(6017): 589-592.

Yeates, T. O., J. Jorda and T. A. Bobik. 2013. "The shells of BMC-type microcompartment organelles in bacteria." *J Mol Microbiol Biotechnol* 23(4-5): 290-299.

Zhu, X. D. and P. D. Sadowski. 1995. "Cleavage-dependent ligation by the FLP recombinase. Characterization of a mutant FLP protein with an alteration in a catalytic amino acid." *J Biol Chem* 270(39): 23044-23054.

Chapter 6

Conclusions and Perspectives

Engineering carboxysomes in heterologous hosts for new functions has attracted increasing attention recently. Two types of carboxysomes consisting of similar structural morphology show great potentials of bioengineering. In this work, we conducted the synthetic engineering of α- and β-carboxysomes as well as empty carboxysome shells in *E. coli* and characterized the composition, structure, and function of the produced synthetic organelles. In Chapter 3, we expressed β-carboxysomes from *Synechococcus elongatus* PCC 7942 into *E. coli*. The plasmids contain the genes encoding β-carboxysome proteins were constructed. Expression of β-carboxysome proteins in BL21(DE3) cells were determined by SDS-PAGE, western blotting, and mass spectrometry, though at a low level. The generated carboxysome-like structures were observed by EM and AFM, and the activity of synthetic carboxysomes was examined by Rubisco assays *in vivo* and *in vitro*. Furthermore, the interchangeability of proteins from different BMC origins was characterized. In Chapter 4, synthetic α-carboxysomes were produced in *E. coli* and their assembly and Rubisco activity were described in detail. We further redesigned the α-carboxysome-expressing constructs by adding the gene encoding Rubisco activase CbbQ. This resulted in an enhanced Rubisco activity and no significant effects on the composition and assembly of synthetic α-carboxysomes. In Chapter 5, intact empty shells (50~80 nm in diameter) based on the α-carboxysome structure were generated using the *E. coli* expression system and were isolated at a high yield for structural and functional characterization. Producing and modulating intact carboxysome shells lay the foundation for constructing new nanobioreactors and

nanocages, based on BMC structures, with the potentials of adjusting metabolic activities.

In Chapter 3, the low-level protein expression and unspecific shape with limited Rubisco activity of synthetic β-carboxysomes reveal the necessity of understanding the native carboxysome assembly and function for modifying the protein expression and assembly conditions to improve carboxysome engineering. For instance, the exact abundance of individual components of α- and β-carboxysomes and their regulation at the transcript and translation levels which are still a mystery. Moreover, it is vital to characterize the proteins/factors involved in the assembly and function of carboxysomes. Heterogeneous engineering of active Rubisco requires several proteins, such as Rubisco activase and chaperones, which include chaperonins Cpn60/Cpn20, RuBisCo accumulation factors 1 and 2, RbcX, and bundle-sheath defective-2 (BSD2) (Aigner et al., 2017). Because of the assembly of β-carboxysome initiates from Rubisco core aggregation, these Rubisco assembly factors might be introduced to improve the structure and function of synthetic carboxysomes. In this study, we tested the effects of the Rubisco activase CbbQ on the activity of synthetic α-carboxysome. The future study will focus on the roles of CbbO in engineering α-carboxysome structures for enhanced Rubisco activity.

Engineering the cyanobacterial CCM, especially carboxysomes into the chloroplasts is a convincing approach to improving plant photosynthesis. However, previous research only showed transient expression of β-carboxysomes shell proteins could form aggregations in chloroplasts (Lin et al., 2014), and chloroplasts transformation is

still arduous in most of the crop plants. So studying the engineering of carboxysome in *E. coli* can help us to understand the possibility of the stable carboxysome assembly in plant and further simplify this process. However, depending on my results, the β-carboxysome structures synthesized in *E. coli* are imperfect, which means synthesized functional β-carboxysome in plants might not achievable for now. The α-carboxysome and α-empty shell can be self-assembled in *E. coli* which exhibited the possibility of α-carboxysome assembly in the plant. However, the different surface charges of Rubisco Form 1A and 1B (Zarzycki et al., 2013) indicated inserting plant Form 1B Rubisco into synthetic α-carboxysomes may not be possible. But we still can try whether Form 1A Rubisco can work appropriately in plant or capture Form 1B Rubisco into α-carboxysome shells with the help of EP.

6.1 Reference

Aigner, H., R. H. Wilson, A. Bracher, L. Calisse, J. Y. Bhat, F. U. Hartl and M. Hayer-Hartl. 2017. "Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2." *Science* 358(6368): 1272-1278.

Lin, M. T., A. Occhialini, P. J. Andralojc, J. Devonshire, K. M. Hines, M. A. Parry and M. R. Hanson. 2014. "beta-Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts." *Plant J* 79(1): 1-12.

Zarzycki, J., S. D. Axen, J. N. Kinney and C. A. Kerfeld. 2013. "Cyanobacterial-based approaches to improving photosynthesis in plants." *J Exp Bot* 64(3): 787-798.